```
   1 GAACCCAGTT GCTTCAGCGA GTCGAACTAC AGTTTTAACC TCATCAAATA
  51 TGGCATCTCC CTTGCTTGCT GCAGCAGGGA TGGAAGAAAT GTCACTTTCT
 101 TTTTAAGCTA GCAAGCTTTT TCTTTTTCTT TTTCTTCTTC TATTTAAAAA
 151 TTCTAATCAT GGATGCTTCT TCCGACCCTT ATTTGCCTTA TGACGGGGGA
 201 GGAGACAATA TTCCCCTGAG GGAATTACAT AAAAGAGGAA CTCATTATAC
 251 AATGACAAAT GGAGGCAGCA TTAACAGTTC TACACATTTA CTGGATCTTT
 301 TGGATGAACC AATTCCAGGT GTTGGTACAT ATGATGATTT CCATACTATT
 351 GATTGGGTGC GAGAAAAATG TAAAGACAGA GAAAGGCATA GACGGATCAA
 401 CAGCAAAAAG AAAGAATCAG CATGGGAAAT GACAAAAAGT TTGTATGATG
 451 CGTGGTCAGG ATGGCTAGTA GTAACACTAA CAGGATTGGC ATCAGGGGCA
 501 CTGGCCGGAT TAATAGACAT TGCTGCCGAT TGGATGACTG ACCTAAAGGA
 551 GGGCATTTGC CTTAGTGCGT TGTGGTACAA CCACGAACAG TGCTGTTGGG
 601 GATCTAATGA AACAACATTT GAAGAGAGGG ATAAATGTCC ACAGTGGAAA
 651 ACATGGGCAG AATTAATCAT AGGTCAAGCA GAGGGTCCTG GTTCTTATAT
 701 CATGAACTAC ATAATGTACA TCTTCTGGGC CTTGAGTTTT GCCTTTCTTG
 751 CAGTTTCCCT GGTAAAGGTA TTTGCTCCAT ATGCCTGTGG CTCTGGAATT
 801 CCAGAGATTA AAACTATTTT AAGTGGATTC ATCATCAGAG GTTACTTGGG
 851 AAAATGGACT TTAATGATTA AAACCATCAC ATTAGTCCTG GCTGTGGCAT
 901 CAGGTTTGAG TTTAGGAAAA GAAGGTCCCC TGGTACATGT TGCCTGTTGC
 951 TGCGGAAATA TCTTTTCCTA CCTCTTTCCA AAGTATAGCA CAAACGAAGC
1001 TAAAAAAGG GAGGTGCTAT CAGCTGCCTC AGCTGCAGGG GTTTCTGTAG
1051 CTTTTGGTGC ACCAATTGGA GGAGTTCTTT TTAGCCTGGA AGAGGTTAGC
1101 TATTATTTTC CTCTCAAAAC TTTATGGAGA TCATTTTTTG CTGCTTTAGT
1151 GGCTGCATTT GTTTTGAGGT CCATCAATCC ATTTGGTAAC AGCCGTCTGG
1201 TCCTTTTTTA TGTGGAGTAT CATACACCAT GGTACCTTTT TGAACTGTTT
1251 CCTTTTATTC TTCTAGGGGT ATTTGGAGGG CTTTGGGGAG CCTTTTTCAT
1301 TAGGGCAAAT ATTGCCTGGT GTCGTCGACG CAAGTCCACG AAATTTGGAA
1351 AGTATCCCGT TCTGGAAGTC ATTATTGTTG CAGCCATTAC TGCTGTGATA
1401 GCCTTCCCTA ATCCATACAC TAGGCTAAAC ACCAGTGAAC TGATCAAAGA
1451 GCTTTTTACA GACTGTGGTC CCCTGGAATC CTCTTCTCTT TGTGACTACA
1501 GAAATGACAT GAATGCCAGT AAAATTGTCG ATGACATTCC TGATCGTCCA
1551 GCAGGCATTG GAGTATATTC AGCTATATGG CAGTTATGCC TGGCACTCAT
1601 ATTTAAAATC ATAATGACAG TATTCACTTT TGGCATCAAG GTTCCATCAG
1651 GCTTGTTCAT CCCCAGCATG GCCATTGGAG CGATCGCAGG AAGGATTGTG
1701 GGGATTGCGG TGGAGCAGCT TGCCTACTAT CACCACGACT GGTTTATCTT
1751 TAAGGAGTGG TGTGAGGTCG GGGCTGATTG CATTACACCT GGCCTTTATG
1801 CCATGGTTGG TGCTGCTGCA TGCTTAGGTG GTGTGACAAG AATGACTGTC
1851 TCCCTGGTGG TTATTGTTTT TGAGCTTACT GGAGGCTTGG AATATATTGT
1901 TCCCCTTATG GCTGCAGTCA TGACCAGTAA ATGGGTTGGA GATGCCTTTG
1951 GCAGGGAAGG CATTTATGAA GCACACATCC GATTAAATGG ATACCCTTTC
2001 TTGGATGCAA AAGAAGAATT CACTCATACC ACCCTGGCTG CTGACGTTAT
2051 GAGACCTCGA AGGAATGATC CTCCCTTAGC TGTCCTGACA CAGGACAATA
2101 TGACAGTGGA TGATATAGAA AACATGATTA ATGAAACCAG CTACAATGGA
2151 TTTCCTGTCA TAATGTCAAA AGAATCTCAG AGATTAGTGG GATTTGCCCT
2201 CAGAAGAGAC CTGACAATTG CAATAGAAAG TGCCAGGAAA AAACAAGAAG
2251 GTATCGTTGG CAGTTCTCGG GTGTGTTTTG CACAGCACAC CCCATCTCTT
2301 CCAGCAGAAA GTCCTCGGCC ATTGAAGCTT CGAAGCATTC TTGACATGAG
2351 CCCTTTTACA GTGACAGACC ACACCCCAAT GGAGATTGTG GTGGATATTT
2401 TCCGAAAGCT GGGACTGAGG CAGTGCCTTG TAACTCACAA TGGGCGCCTC
2451 CTTGGCATTA TAACAAAAAA AGATATCCTC CGGCATATGG CCCAGACGGC
2501 AAACCAAGAC CCCGCTTCAA TAATGTTCAA CTGAATCTCA CAGATGAGGA
2551 GAGAGAAGAA ACGGAAGAGG AAGTTTATTT GTTGAATAGC ACAACTCTTT
2601 AACCTGAGGG AGTCATCTAC TTTTTTTTCC TCCTTTACAA AAAAAGAAAG
2651 GAAATATAAA AGCCGGGTTT TTGCAACATG GTTTGCAAAT AATGCTGGTG
2701 GAATGGAGGA GTTGTTTGGG GAGGGAAAGG AGAGAGAAGG AAAGGAGTGA
2751 GGTATTTCCC GTCTAACAGA AAGCAGCGTA TCAACTCCTA TTGTTCTGCA
2801 CTGGATGCAT TCAGCTGAGG ATGTGCCTGA TAGTGCAGGC TTGCGCCTCA
2851 ACAGAGATGA CAGCAGAGTC CTCGAGCACC TGGCCTGTTG CTCCAACATT
2901 GCAAAGACAC ATTATCAGTC CCTATTTCTA GAGGGATTAC TTTGAATTGA
2951 GCCATCTATA AAACTGCAAG GTCTTGCCCT TTTTTTTAAT CAAAACTGTT
3001 CTGTTTAATT CATGAATTGT ATAGTTAAGC ATTACCTTTC TACATTCCAG
3051 AAGAGCCTTT ATTTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTACTG
3101 AGCTGTAACA AAGCCTCTTT AAATCGGTGT ATCCTTTTGA AGCAGTCCTT
```

```
3151 TCTCATATTG AGATGTACTG TGATTTTACT GAGGTTTCAT CACAAGAAGG
3201 GAGTGTTTCT TGTGCCATTA ACCATGTAGT TTGTACCATC ACTAAATGCT
3251 TGGAACAGTA CACATGCACC ACAACAAAGG CTCATCAAAC AGGTAAAGTC
3301 TCGAAGGAAG CGAGAACGAA ATCTCTCATT GTGTGCCGTG TGGCTCAAAA
3351 CCGAAAACAA TGAAGCTTGG TTTTAAAGGA TAAAGTTTTC TTTTTTGTTT
3401 TCCTCTCAGA CTTTATGGAT AATGTGACCG GGTCTTATGC AAATTTTCTA
3451 TTTCTAAAAC TACTACTATG ATATACAAGT GCTGTTGAGC ATAATTAAAT
3501 AAAATGCTGC TGCTTTGACA GTAAAGAGAA AAAAAAAAAA AAAAAAAAAA
3551 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
3601 AAAAAAAAAA AAAAAAAAAA AAAAA  (SEQ ID NO:1)
```

FEATURES:
5'UTR: 1-158
Start Codon: 159
Stop Codon: 2532
3'UTR: 2535

FIGURE 1, page 2 of 3

HOMOLOGOUS PROTEINS:

Top BLAST Hits:

	Score	E
CRA|18000005109762 /altid=gi|2599548 /def=gb|AAB95161.1| (AF029...	1575	0.0
CRA|18000005109763 /altid=gi|2599550 /def=gb|AAB95162.1| (AF029...	1573	0.0
CRA|18000005227216 /altid=gi|4762023 /def=gb|AAD29440.1|AF14277...	1572	0.0
CRA|18000004989660 /altid=gi|4502869 /def=ref|NP_001820.1| chlo...	1570	0.0
CRA|18000005231972 /altid=gi|8134363 /def=sp|Q9R279|CLC3_CAVPO ...	1561	0.0
CRA|18000004989700 /altid=gi|6680948 /def=ref|NP_031737.1| chlo...	1560	0.0
CRA|18000004978791 /altid=gi|1705905 /def=sp|P51792|CLC3_RAT CH...	1560	0.0
CRA|1000685681515 /altid=gi|6634696 /def=emb|CAA71072.2| (Y0994...	1559	0.0
CRA|18000004989661 /altid=gi|1705903 /def=sp|P51790|CLC3_HUMAN ...	1558	0.0
CRA|18000005226296 /altid=gi|4753144 /def=gb|AAB88634.2| (U8346...	1556	0.0

EST:

	Score	E
gi|10993825 /dataset=dbest /taxon=96...	1562	0.0
gi|10934924 /dataset=dbest /taxon=96...	1336	0.0
gi|10952244 /dataset=dbest /taxon=96...	1251	0.0
gi|12383593 /dataset=dbest /taxon=96...	1205	0.0
gi|6591096 /dataset=dbest /taxon=9606 ...	1170	0.0
gi|10251711 /dataset=dbest /taxon=96...	1104	0.0
gi|2321385 /dataset=dbest /taxon=9606 ...	1045	0.0
gi|5594360 /dataset=dbest /taxon=9606 ...	975	0.0
gi|5422132 /dataset=dbest /taxon=9606 ...	965	0.0
gi|10327969 /dataset=dbest /taxon=96...	963	0.0

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

gi|10993825	Neuronal precursor cells-teratocarcinoma
gi|10934924	Whole embryo-mainly head
gi|10952244	Neuronal precursor cells-teratocarcinoma
gi|12383593	Small intestine-duodenal adenocarcinoma
gi|6591096	Lung-small cell carcinoma
gi|10251711	Breast-normal
gi|2321385	Schwannoma tumor
gi|5594360	Brain-tumor
gi|5422132	Testis
gi|10327969	Lung-large cell carcinoma

FIGURE 1, page 3 of 3

```
  1 MDASSDPYLP YDGGGDNIPL RELHKRGTHY TMTNGGSINS STHLLDLLDE
 51 PIPGVGTYDD FHTIDWVREK CKDRERHRRI NSKKKESAWE MTKSLYDAWS
101 GWLVVTLTGL ASGALAGLID IAADWMTDLK EGICLSALWY NHEQCCWGSN
151 ETTFEERDKC PQWKTWAELI IGQAEGPGSY IMNYIMYIFW ALSFAFLAVS
201 LVKVFAPYAC GSGIPEIKTI LSGFIIRGYL GKWTLMIKTI TLVLAVASGL
251 SLGKEGPLVH VACCCGNIFS YLFPKYSTNE AKKREVLSAA SAAGVSVAFG
301 APIGGVLFSL EEVSYYFPLK TLWRSFFAAL VAAFVLRSIN PFGNSRLVLF
351 YVEYHTPWYL FELFPFILLG VFGGLWGAFF IRANIAWCRR RKSTKFGKYP
401 VLEVIIVAAI TAVIAFPNPY TRLNTSELIK ELFTDCGPLE SSSLCDYRND
451 MNASKIVDDI PDRPAGIGVY SAIWQLCLAL IFKIIMTVFT FGIKVPSGLF
501 IPSMAIGAIA GRIVGIAVEQ LAYYHHDWFI FKEWCEVGAD CITPGLYAMV
551 GAAACLGGVT RMTVSLVVIV FELTGGLEYI VPLMAAVMTS KWVGDAFGRE
601 GIYEAHIRLN GYPFLDAKEE FTHTTLAADV MRPRRNDPPL AVLTQDNMTV
651 DDIENMINET SYNGFPVIMS KESQRLVGFA LRRDLTIAIE SARKKQEGIV
701 GSSRVCFAQH TPSLPAESPR PLKLRSILDM SPFTVTDHTP MEIVVDIFRK
751 LGLRQCLVTH NGRLLGIITK KDILRHMAQT ANQDPASIMF N   (SEQ ID NO:2)
```

FEATURES:
Functional domains and key regions:
[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

Number of matches: 5
```
      1        90-93 NETT
      2      364-367 NTSE
      3      392-395 NASK
      4      587-590 NMTV
      5      598-601 NETS
```

[2] PDOC00004 PS00004 CAMP_PHOSPHO_SITE
cAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 3
```
      1        24-27 KKES
      2      330-333 RRKS
      3      331-334 RKST
```

[3] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 8
```
      1        22-24 SKK
      2      333-335 STK
      3      529-531 TSK
      4      613-615 SQR
      5      631-633 SAR
      6      642-644 SSR
      7      658-660 SPR
      8      709-711 TKK
```

FIGURE 2, page 1 of 5

```
[4] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 13
        1       27-30 SAWE
        2       34-37 SLYD
        3       92-95 TTFE
        4       93-96 TFEE
        5      105-108 TWAE
        6      217-220 STNE
        7      249-252 SLEE
        8      383-386 SLCD
        9      589-592 TVDD
       10      666-669 SILD
       11      674-677 TVTD
       12      679-682 TPME
       13      709-712 TKKD


[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 18
        1       49-54 GLASGA
        2       53-58 GALAGL
        3       72-77 GICLSA
        4       88-93 GSNETT
        5      189-194 GLSLGK
        6      206-211 GNIFSY
        7      234-239 GVSVAF
        8      240-245 GAPIGG
        9      245-250 GVLFSL
       10      310-315 GVFGGL
       11      313-318 GGLWGA
       12      314-319 GLWGAF
       13      408-413 GVYSAI
       14      447-452 GAIAGR
       15      491-496 GAAACL
       16      541-546 GIYEAH
       17      638-643 GIVGSS
       18      692-697 GLRQCL
```

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	99	119	1.810	Certain
2	182	202	2.131	Certain
3	233	253	1.398	Certain
4	256	276	1.019	Certain
5	290	310	1.770	Certain
6	321	341	0.797	Putative
7	361	381	2.093	Certain
8	400	420	1.539	Certain
9	473	493	1.739	Certain
10	496	516	1.218	Certain
11	540	560	1.568	Certain
12	570	590	0.975	Putative

FIGURE 2, page 2 of 5

BLAST Alignment to Top Hit:

```
>CRA|18000005109762 /altid=gi|2599548 /def=gb|AAB95161.1| (AF029346)
            chloride channel protein 3 [Homo sapiens] /org=Homo
            sapiens /taxon=9606 /dataset=nraa /length=818
          Length = 818


 Score = 1572 bits (4026), Expect = 0.0
 Identities = 764/765 (99%), Positives = 764/765 (99%)


Query: 27   GTHYTMTNGGSINSSTHLLDLLDEPIPGVGTYDDFHTIDWVREKCKDRERHRRINSKKKE 86
            GTHYTMTNGGSINSSTHLLDLLDEPIPGVGTYDDFHTIDWVREKCKDRERHRRINSKKKE
Sbjct: 54   GTHYTMTNGGSINSSTHLLDLLDEPIPGVGTYDDFHTIDWVREKCKDRERHRRINSKKKE 113


Query: 87   SAWEMTKSLYDAWSGWLVVTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC 146
            SAWEMTKSLYDAWSGWLVVTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC
Sbjct: 114  SAWEMTKSLYDAWSGWLVVTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC 173


Query: 147  WGSNETTFEERDKCPQWKTWAELIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 206
            WGSNETTFEERDKCPQWKTWAELIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA
Sbjct: 174  WGSNETTFEERDKCPQWKTWAELIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 233


Query: 207  PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG 266
            PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG
Sbjct: 234  PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG 293


Query: 267  NIFSYLFPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLEEVSYYFPLKTLWRSF 326
            NIFSYLFPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLEEVSYYFPLKTLWRSF
Sbjct: 294  NIFSYLFPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLEEVSYYFPLKTLWRSF 353


Query: 327  FAALVAAFVLRSINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGAFFIRANIA 386
            FAALVAAFVLRSINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGAFFIRANIA
Sbjct: 354  FAALVAAFVLRSINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGAFFIRANIA 413


Query: 387  WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD 446
            WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD
Sbjct: 414  WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD 473


Query: 447  YRNDMNASKIVDDIPDRPAGIGVYSAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMAI 506
            YRNDMNASKIVDDIPDRPAGIGVYSAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMAI
Sbjct: 474  YRNDMNASKIVDDIPDRPAGIGVYSAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMAI 533


Query: 507  GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL 566
            GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL
Sbjct: 534  GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL 593


Query: 567  VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHTTL 626
            VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHTTL
Sbjct: 594  VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFTHTTL 653


Query: 627  AADVMRPLRNDPPLAVLTQDNMTVDDIENMINETSYNGFPVIMSKESQRLVGFALRRDLT 686
            AADVMRP RNDPPLAVLTQDNMTVDDIENMINETSYNGFPVIMSKESQRLVGFALRRDLT
Sbjct: 654  AADVMRPRRNDPPLAVLTQDNMTVDDIENMINETSYNGFPVIMSKESQRLVGFALRRDLT 713


Query: 687  IAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLRSILDMSPFTVTDHTPMEIVVD 746
            IAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLRSILDMSPFTVTDHTPMEIVVD
Sbjct: 714  IAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLRSILDMSPFTVTDHTPMEIVVD 773


Query: 747  IFRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 791
            IFRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN
Sbjct: 774  IFRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 818    (SEQ ID NO:4)
```

FIGURE 2, page 3 of 5

```
>CRA|18000004989660 /altid=gi|4502869 /def=ref|NP_001820.1| chloride
        channel 3; ClC-3 [Homo sapiens] /org=Homo sapiens
        /taxon=9606 /dataset=nraa /length=820
        Length = 820

 Score = 1567 bits (4013), Expect = 0.0
 Identities = 764/767 (99%), Positives = 764/767 (99%), Gaps = 2/767 (0%)

Query: 27   GTHYTMTNGGSINSSTHLLDLLDEPIPGVGTYDDFHTIDWVREKCKDRERHRRINSKKKE 86
            GTHYTMTNGGSINSSTHLLDLLDEPIPGVGTYDDFHTIDWVREKCKDRERHRRINSKKKE
Sbjct: 54   GTHYTMTNGGSINSSTHLLDLLDEPIPGVGTYDDFHTIDWVREKCKDRERHRRINSKKKE 113

Query: 87   SAWEMTKSLYDAWSGWLVVTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC 146
            SAWEMTKSLYDAWSGWLVVTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC
Sbjct: 114  SAWEMTKSLYDAWSGWLVVTLTGLASGALAGLIDIAADWMTDLKEGICLSALWYNHEQCC 173

Query: 147  WGSNETTFEERDKCPQWKTWAELIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 206
            WGSNETTFEERDKCPQWKTWAELIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA
Sbjct: 174  WGSNETTFEERDKCPQWKTWAELIIGQAEGPGSYIMNYIMYIFWALSFAFLAVSLVKVFA 233

Query: 207  PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG 266
            PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG
Sbjct: 234  PYACGSGIPEIKTILSGFIIRGYLGKWTLMIKTITLVLAVASGLSLGKEGPLVHVACCCG 293

Query: 267  NIFSYLFPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLEEVSYYFPLKTLWRSF 326
            NIFSYLFPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLEEVSYYFPLKTLWRSF
Sbjct: 294  NIFSYLFPKYSTNEAKKREVLSAASAAGVSVAFGAPIGGVLFSLEEVSYYFPLKTLWRSF 353

Query: 327  FAALVAAFVLRSINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGAFFIRANIA 386
            FAALVAAFVLRSINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGAFFIRANIA
Sbjct: 354  FAALVAAFVLRSINPFGNSRLVLFYVEYHTPWYLFELFPFILLGVFGGLWGAFFIRANIA 413

Query: 387  WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD 446
            WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD
Sbjct: 414  WCRRRKSTKFGKYPVLEVIIVAAITAVIAFPNPYTRLNTSELIKELFTDCGPLESSSLCD 473

Query: 447  YRNDMNASKIVDDIPDRPAGIGVYSAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMAI 506
            YRNDMNASKIVDDIPDRPAGIGVYSAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMAI
Sbjct: 474  YRNDMNASKIVDDIPDRPAGIGVYSAIWQLCLALIFKIIMTVFTFGIKVPSGLFIPSMAI 533

Query: 507  GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL 566
            GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL
Sbjct: 534  GAIAGRIVGIAVEQLAYYHHDWFIFKEWCEVGADCITPGLYAMVGAAACLGGVTRMTVSL 593

Query: 567  VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKE--EFTHT 624
            VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKE  EFTHT
Sbjct: 594  VVIVFELTGGLEYIVPLMAAVMTSKWVGDAFGREGIYEAHIRLNGYPFLDAKEEFEFTHT 653

Query: 625  TLAADVMRPLRNDPPLAVLTQDNMTVDDIENMINETSYNGFPVIMSKESQRLVGFALRRD 684
            TLAADVMRP RNDPPLAVLTQDNMTVDDIENMINETSYNGFPVIMSKESQRLVGFALRRD
Sbjct: 654  TLAADVMRPRRNDPPLAVLTQDNMTVDDIENMINETSYNGFPVIMSKESQRLVGFALRRD 713

Query: 685  LTIAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLRSILDMSPFTVTDHTPMEIV 744
            LTIAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLRSILDMSPFTVTDHTPMEIV
Sbjct: 714  LTIAIESARKKQEGIVGSSRVCFAQHTPSLPAESPRPLKLRSILDMSPFTVTDHTPMEIV 773

Query: 745  VDIFRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 791
            VDIFRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN
Sbjct: 774  VDIFRKLGLRQCLVTHNGRLLGIITKKDILRHMAQTANQDPASIMFN 820   (SEQ ID NO:5)
```

```
>CRA|1000685681515 /altid=gi|6634696 /def=emb|CAA71072.2| (Y09941)
            putative chloride channel ClC-3 [Xenopus laevis]
            /org=Xenopus laevis /taxon=8355 /dataset=nraa
            /length=791
          Length = 791


 Score = 1559 bits (3993), Expect = 0.0
 Identities = 745/791 (94%), Positives = 771/791 (97%)


Query: 1    MDASSDPYLPYDGGGDNIPLRELHKRGTHYTMTNGGSINSSTHLLDLLDEPIPGVGTYDD 60
            MD SSDPYLPYDGGGDNIPLR+LHKRGTHYT+TNGG+INS+THLLDLLDEPIPGVGTYDD
Sbjct: 1    MDISSDPYLPYDGGGDNIPLRDLHKRGTHYTVTNGGAINSTTHLLDLLDEPIPGVGTYDD 60
(SEQ ID NO:6)



**Hmmer search results (Pfam):**
Model       Description                                        Score      E-value   N
CE00039     CE00039 chloride_channel                           1671.9          0    1
CE00420     E00420 CLC                                         1288.1          0    2
PF00654     Voltage gated chloride channels                    1172.4          0    1
PF00571     CBS domain                                           78.1      7e-20    2


Parsed for domains:
Model       Domain    seq-f seq-t     hmm-f hmm-t       score   E-value
PF00654     1/1          71   622 ..      1   621 []    1172.4        0
PF00571     1/2         645   690 ..     11    54 .]      31.4  5.8e-07
CE00420     1/2          32   697 ..      1   729 [.    1174.4        0
PF00571     2/2         726   778 ..      1    54 []      47.4  2.2e-11
CE00420     2/2         722   791 .]    867   942 ..     110.6  6.5e-32
CE00039     1/1          60   791 .]      1   804 []    1671.9        0
```

<center>FIGURE 2, page 5 of 5</center>

```
   1 AATTCTATAC AAATATAATT ATATAGATAT ATATTACATA TACACACAAT
  51 TGTTTATCTT TAAAAATAAT TCAAATATGG CTACAAAACT TTTACAATAT
 101 GAAGCATTGT CAGTATTTAT TTTACCGGGA GGATTCCCC CATCAGTGAG
 151 TGCTGACTGT CATTTTCATT CTTTATGATC AAGTTGTAGA TCAGGAAAAA
 201 CAAGTTAAGA GAGTGCCTAC AAATACCGGG AAAACTTGTG GATAGATTTT
 251 CATTTTTTAT GTAAAGACAT ATAAGAACAT GAATGGTATA AAAACAAAAT
 301 CCTTTATAAA TGCCATACAA TTATATATTT AGAAAAATTA TATGGTGGTA
 351 AAACATATAA AAGAACCACA CACTCCCAAA TTTACATTGA GCTAATTTAG
 401 TACAGTTAGC CTTTGTCAAA GCTTTCCTTG TTTAAAAAAA CTATTGGCTC
 451 AGTGTGCAGG AAGGAGCATA GGAGAAAAAA TTGCCAAGAA TATTTGAAAA
 501 ATACAGAAAA TAAAGAAAAA AATCACCTAC TATCCTATCA AAAATTTTAA
 551 TAGCTAGAAT CAGGATAAGA TAGAATATTC CTGTGGCAGT AATTCTAGTC
 601 TATATTCCTT TCCTGGAACC CTGTCTCCCA AATTTCAGGT GAGATTTTAT
 651 AAGAAGCTCT GTTTATCTGA GATTTAAAAT ATAAAACTTT GATTTAACCT
 701 ATACAGTTTT TTAAAAAGAC CCTAAATAAG TAAAATTTAG TACTCCACAA
 751 ATTGAAGAGA ATTTCTCTCT TCTCTTTACT GCCCTCTGAG TTTTCTCTTT
 801 CCTTCTCTCA CCTCCAATTT TCATGTAAAC ACTTTCAGTT CGAGTGGACC
 851 TTAGAGATTG TCTCATTCAA TACTTTAGGA AAACAAATTT TATAGAACCC
 901 TTGAGTTCTG TGGAATTGCT TCTAATGAAC AACACCTTTT GTTGTTGTTG
 951 TTGTTTAGTG ACACTGTGTA ACAGGCATTT CAGGAGGAGA ATCTCCCAGT
1001 CTAGAGGAAT CCTCTCAGAG GTAGCTATAA AATATTGAAC TCTGATCTTC
1051 AATAAGCATT GTGCGGTTTT TGTTTTTGTT TTAATGACA GTTTTAAACA
1101 AGAAAGTTGC TTTATTTCTG AACTTCATAA AAATTTCTAT TAAAGAGACA
1151 ATTTCTGAAT TTTATAACAA TTTCTAGAAC AGTTGAGTAC CTCACTTTGA
1201 GACACATTTT TGCTAAAAGT TAAAAACACA AAACCCTTAT GAGATAAAAT
1251 AGGAAGCTAG TAGAGATAGG AAAGTCCTCT GCTTAGTAAA CCTCTTTTTT
1301 GCGTAGTTTA GACACATACA ATAGTAAAGT TACTTAGTAC GTTGATAGTT
1351 TTCTTTCTCC TCAAAAGCTA CAATGTCTTA CTAGCTAGTT CCTTCAAGAA
1401 AGGAAACAAG AAGCCGCTGG AGGAGATTGG TGAGTGGGAT AAAACACTAT
1451 TCAACTCTTC AGTTATTCGG TTTTTAAATC CTCAATGAAA GGCTGCTGTA
1501 TTATAGAGTA TTTTTTTTTT ATTTTTAATA GACTTAGAAC CAAGTTTCTT
1551 GAGAAACCTT TGGCATATTG TAGTTTTTTT ATGGCTATGA CTCACATGAC
1601 ATTACTGTAT AAAACTAGTA CATTCTCTCG TAAAACCACA CAAACTTACT
1651 AGAGTGCTGC TCTCATTTTT CTACATTAGA AATGAAAAAG GCATTGTCT
1701 GCATTCAAAA TTTCCTTTTT ACATCTCTGT ATTACTTTTT CCCCTTTATA
1751 TTTATCTTAA AACCAAAAGA AATAATGTTT CTATTGTTTT ACTGTAGTTA
1801 CCACTGATGC TACCGAAGCT GTATTGTGAG TGTTTCAAAA TTCTCAAACC
1851 AGTTTTGTGT GTTGTACTTG GAGCTTAGTC ATTGTCATAC GTAGCAGGAC
1901 CTGATTAAGA AGGCTGTGCC GCCTCTAAGC CTTGCTAGAT TGTAGCCACT
1951 AGCAACCAGG CTGCAATAAT TTCCCTTTGA TGACATCATC CACTGTGGAA
2001 GAACCCAGTT GCTTCAGCGA GTCGAACTAC AGTTTTAACC TCATCAAATA
2051 TGGCATCTCC CTTGCTTGCT GCAGCAGGGA TGGAAGAAAT GTCACTTTCT
2101 TTTTAAGCTA GCAAGCTTTT TCTTTTTCTT TTTCTTCTTC TATTTAAAAA
2151 TTCTAATCAT GGATGCTTCT TCCGACCCTT ATTTGCCTTA TGACGGGGGA
2201 GGAGACAATA TTCCCCTGAG GGAATTACAT AAAAGAGGTA ATACTATCCC
2251 CTTGCTGTGA ATTCTCTGTT GGTATGTTTT GCATGCGGCT GGGCGGTCCT
2301 CTAGCTTAAA CTGGTTCTCG TTTGTCCTTT AAATACTGCA GTACGTTGTT
2351 TAGTTGCCCT GGGTTGTTAG TAAGGGGAAA ATGCAACCTT CTGAATGGTT
2401 GTGTAGCCAT CCCTGATTGT TTTCTCTGTG CAGATTAGTA CTGCTTCAGA
2451 TCACGTCGGG CTCCGACTCC ATCTTCTGCA TGAAAATCTT CTTTCTAACT
2501 CTGAAAATGA ATTAATCTGC TTTTACAGCC AACTAAAGTC GTGTTGGTTG
2551 GCATCTAAAA AGTAATGTTT TTCTTCCTTC AGAAAACTTA CATTTCCTTT
2601 AATTTACACA GAGAAATCAG GTGCCTATGT ACCATTATAT TTTAGCTGCT
2651 GCCAATTACC ATGTAGATTT TACACCACAA AGTAAATTTA TAGCAAAAGC
2701 TTTACCTACA TTTTAGAACA TTTTAAAATG ATAGTAAAGA TGAATAATTT
2751 CTATATTAAT ACTTTTTATT TAATATGTAT TTCGGCTGAG TAACATACTA
2801 CATTGTCTTC CACAGGTATC TTGTGAAATT TGATATGATA AAACACATTT
2851 GACTAAATGT CAGAAAAAAT AATATTGGTT TGTGAAAAGC AGAAGAGCAC
2901 CCAGCATGCC TGTAAATCTT TTGGCAGGCA CTTCCTCAGT CTCCTTAAAA
2951 TTAATTGCAT GTTAATTACT ACCCTTTTTT TCATTTTTGT TTAATTGCTT
3001 ATTCGAAAAA CAGACTGGTC GACATTTGTT GTCCTAGAAA AAAATTGAAC
3051 TTCAAGAAAA ATCTCTTAGC TTATGTGACT TCATTTTTGA GCCACATTAG
3101 TTTGAATTAC TGCATGATAT TATAAACTCA CCTTATGATT TAACCCAAAC
```

FIGURE 3, page 1 of 27

```
3151 TTTTATTTGT AAGTATATAA GGAAGTAATA ATGTTTTTCT AATATAATTA
3201 GCCTGCTTTA TTTAAAATAT ACTTTGTGTT CTGATAACAC TTTTTTTTTA
3251 GTATTAAGTT CCACTATAAT TTAAACATTA TAATGTATTC AACAAATGTC
3301 TGTTGGTTGC ATTGTGTCTG CTACACACTA TTTTAGGGTC TGAACAGTTG
3351 TAGCATTATT TATCTTGCAG TATTCTGTAG TTAGTAAAAA CTTGCTTTTT
3401 ACATTTTGAG AAAAGCTGTG TAAGGATCAT GTTACATACA TTGTGCTTTC
3451 TCTTACAGAG TTACCTTCTT AATAAAATTT TGATATATGT GTATATGTAT
3501 ATGTTAGAAC ATTTGGAAGA AATATCTAAA AGCATAAAGA AGAAAATAAT
3551 TTCTTGTAAT CACACCACCC AGAGCTTTTT AAATTTTTTT TCTTAATGTT
3601 ACGATCATAA ATTCTTCTAT TTCCTATGTT CTGATTATCA GTTTTCTGGT
3651 AAGGAGTTCT TTAAACAGGA AGCAAGGTGA ATGAATAGTG ACTGTTCAAA
3701 TGTCACATTA TTTGCTAATC AGTAATTAAA CTGTAAAACA AGACAGACTG
3751 TATTTTCCTC ATGCTATTAC AACATTTGGT TGTTAATGAT GATAGATCAG
3801 AATACCTGGG CTTCAGAAAT TTAAATTCCT TTTGTGAAGC TTAACAGTCT
3851 TTGACAGAAC TTACTTATGG ACTGTCTTAG TGTAAAATAT GCAAATAATA
3901 AGAAATAAGT CAAAACTTAT GTGAGAGTAG GCATGGTTAC TGATATTACC
3951 TAAACGTAAG CTTTTTATTT CTATTATACT TTCATAAATA ATCCTTTAAG
4001 AATCTTGCTT AGGATCTAAA TCAGTCCCAC TCTTGGCAGC TCAAATAGGT
4051 TCTTTATCCC TTGATGAGAC TTATTCTATT AATATAAGTC ATTGTTATTT
4101 GAAAGTAACA TTGTGTATGT GTAGTAGAGA TAAGTCAGTT ATTAGGCTTT
4151 CGTGACTGTA CTGTATTACC TCAAACATAC TGTAGTATCC TAGTGTCTAT
4201 GCGTAAGATG TTATTTTTTG TCCATAATTT ATGACCTGTT GTAGCCATGG
4251 GTCAACACAA TGGAATTGAT GGAGACAGGC AGCTAACAAA TCGAAAAAAC
4301 TGAATCAGCT TCCCTGTGAG GAAGAACAAA ACTATAATGA TTAAAATTGA
4351 TCTTCAGCCT GATAGTGAAG AGGCAGATAA AGTATAAAAT TGTGAAGGAT
4401 ATCAATAAAG TAAACATGGA TCTGTTTAGT AAATCCCTGA GTGCTATAGC
4451 CAAGGATTAC CTTTGTTGAG TAAATTGAAT TTAATACTAC TTTTCAAGGC
4501 GAGATGGTAA ATGGTGAAGC TTCCTATTTA AGTAAATAAT GTCAAGTCTG
4551 GAAGTATAAG TAGATTCAAA TTAGAATTAG TTTGATATAC TATTGATAGA
4601 TTAGAAATTA AGATGACATT TCAGAAATAG CCATCTTTAG GGGTAGATTT
4651 CCTATATAGA AACAATCAAG CTCTCTCAAA ATGTCTCTTC CTTTTTTATC
4701 AGGAAAAAAG ACTTGGCTTA TCTGGACTGT TAGTTTTACA CTTTTTTCTTC
4751 TTAATTTGTT CAAGATGTTT AAGTAGTTTT AGAGGTCAAA TTTCTTTCTT
4801 CTACCAACCC TTTATAATGG ATTTGATTCT TTTGGGCCTG AGCCTCCATT
4851 TACTCCATGA GGGGCCTTTA ACAATTATTT AAATNNNNNN NNNNNNNNNN
4901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
4951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
5101 NNNNNNNNNN NNNNNAAAAT AGTAATATTA ATAATAGTTA ATATTTATTA
5151 GAATTTCCTG TTAGCTGGAT ACTGTCCCTA AGTGGGTTTT TTGTTGTTG
5201 TTGTTGTTGT TGTTGTTTTC TTAAGAGAGA GGTATCACTT TTTCACCCAG
5251 GCTGGAGTGC AGTGGAGTGA TTATAGCAAA TGCAGCCTTG AACTACTGGG
5301 CTTAGATCCT CCGTCTCACC CTCCTTGGTA CCTGGGACTG CAGGCTTGCA
5351 ACACCTTGCC TGGCTAATTT AAAAAACAAA ATTTTTTTTT TTTTAGGGA
5401 GAGTCTCACT ATGTTGTCCA GGCTGGTCTC CAACTCCTGG GCTCAAGCAA
5451 CCCTCCTGCC TTGGCCTCCC AAGTAGCTGA GATTACAGGT GCGAGCCACT
5501 GTGCCTGGCT TGTTCTAAGT GCTTTATGTG TATGAAATTA TTTAAATCCT
5551 CATCACAAGT TTATGAAGTA GGTACTGTTA TAATCCCCAT TTTCTAGTTG
5601 ACAAGACTGA GGTAAGGAAT TGTTAAGGAA AAGTCAGAAT TCCATCCAGA
5651 TATTTGGCTC ATACTTTAAT CATGAGGCTA AACTGCTTCT CTCTACACGT
5701 ATCTTCATAG TAACTTGTGT TTTAAGTCTG GTAGAAGCAT AAGAAGTTTA
5751 AACACAGACA GAATCCTGTG GAAGTTAGTA AATTTCTAGT GAACGATAGA
5801 AATGATAGAA ATCTCTTCTT CCCCCAAAGT CCCAAGAACA GATTAGTCTG
5851 CTTTTGACAA GTGTTATCAA AGTAGACTGT TCTCACATAC ACGGGGGACT
5901 CAATAGGGCA TTCCTGGTGG ATATAATAAA ATGAGTAAAT GCGATAACAG
5951 GAGGAAATGC CTAGTGTGTT GCTCTTGGAT TAGTTTTGAT ACAACAAAGG
6001 CAGCTTTGTT GTGAGTCAGT AGAGAGGGTA GTGTAGAACA GTGGAAGTTG
6051 GAAGAGTGGC AGATCCTAGA GGACTAATGA TGGGCTTAAA CCACAAAAAG
6101 TGTCGCTTTG CCATTGAAAT AAAAGTTTGG GGTCTTATTT TTTCAATTTT
6151 CTCCCTGAAA TTATTTCTTG ACATTCATTA GCTCAGCAGT GTATCTAAAT
6201 AAAGCTTTTT TGGGTTTCTA TTATAATAGA GGTTTGTTCC TTTTTCTTCC
6251 CTTTGAAAAG TATCATTTTT TGCACATTAT TTGAAAATCC AGGTGTTATA
```

<div align="center">FIGURE 3, page 2 of 27</div>

```
6301 TGATATTCTT ATTGCCAGAG GGACATTCTG CAGGCTCTTT GTAAAATGAT
6351 TTTAGGATTC AGATACTTAT TATATTTTTA TTGGCCCTAA TATTTTATCC
6401 AACTAGAAAA TTAAACCTCT TCTTAAAAAT TAATCCATCT AAGTGTCTGT
6451 AAATTAAAGG AACAACTAAA GATTCTTTAT TTGGTGTCAG AAACTCCTTG
6501 TTTCTACAAC AGTAGTATAA AACAAAGCCT GTTTTTAAAT GTACTTTTCC
6551 CACAGTATCT GAATTTCAAA TCTTCAATAA AATCTGGTTC ATATTACTAC
6601 CTCTAGCTTG ATTTTCTAAA AATAGCTGAC ACTTTAGTAT GGTTAATTTT
6651 ATGCCATCTC ATGGCTTGTC AGAAATGCTT TGTATCAAGA TTTCCGAGTG
6701 TGAACAGATT TCCTGCCGCA TTGATTAAGT TTGTAATTTT GGCTATTTTC
6751 CCAGCATCGA GGTTTCTGCT TTGCGTTTAT GCAGGAGACT GGTAGTTTAA
6801 ATTGAACTTT AAGGTTTTGT TTCTTGTTTT TAAGTTAACA TATGTTTAAT
6851 TTCTAGTTTC TTTGTAGCCC TTTGCAACTT TAATTAGGTC ATAAAATGGA
6901 TTTACTCTAG TTTCTCTAAC AAATTTTATA AATTTATGAA ATATGAAATT
6951 TAGCAAATTT TATAAACCTT TTTATTCATG TATTGTACAG CTCATCATAT
7001 TTGCAGACAT AATAATTGAA TGTGGAACTT GTTTCCAATT ACACAGATGT
7051 CTTAATATCC ACCTTATCAT CTCTAACTAA AGGATGTGGC TTTTTATTTT
7101 TGAGGTGGCA ACAGAACAGA AAAGAAAACA GTGAATTGAG TAATGGGCTT
7151 AGTATTGCTG CTGCCTGGTT GTGTATCTTT GGTAAACTTC TTTGAGATTT
7201 GGCATTAACT TGCAAGTCTT TGCAGTTTAG ACAGTTAAAT ATGACTGAAT
7251 GGCTGAACAA ATTTTAATAG CGTATGCTTC TTTTTTGCTA TTTATTTACC
7301 CAGTAGACAT TTAATTGACC ACCTGCTAAA TGTGAGGCAC TATTCTTGCC
7351 ATTACCTTTT TAATCTTTGA TTTGGAGTCT GCTAACATTC TGGAACTTCC
7401 ACTATCAACT TAGAACGTTT ACTTTCCCAT CCCTTACCAG GATGGCCATT
7451 TCTTATCAGT AGGGTCACAG AGAGAGAAAA AAAAACCAT CTGGGGCTAG
7501 ACTTCCTGCT CTTAACATAC AGAAGCAAAT AGGTTGTGAA GGAATACATA
7551 GTATTTTGGA TTTCTGCCTC TTCCTTCCAT AATTTTTTTA AAAAGGTTCA
7601 TATGTTTTAT GTGTGTCTTA TGTAACAGTA ATCTGCATTA TGAACTTAAA
7651 TGACGAGGAT CACCATTTCA CATCTTTGGA GATTGATCAC AGAGGTAATA
7701 AGTAACTCTT TTTAAATAAC TATATGCATC ATTTTTCATG TAAAACTATT
7751 ATTTGGATAA ACCCCTTTGA GAAAAGGCTT AGGCTCCTGC CAGTGTCACT
7801 GTGATATTTA CTAATAAGCT CAGTTTAAGG CGCAGCAATT AAGGTTGTGT
7851 TGTTTTTTTT TTTTTAAGTT CAGTTCAGCA AATATATGTG GAAAGCTTGT
7901 GGGTAAAATT ATATTTGTAT TTTTGGGAAA GCAGACAATT TTATTAATGC
7951 CTATATTTTT CTAGTTCAGT GTTTGTCAAA CTTCAAGTTT TAACATGTTG
8001 ATCATGAAAC CAGTTGACTT GTGACCAGTA TTTTAAAAGG AAAGATTAAA
8051 AAAACAAAAT AAAATATCAG TATATACCAA GTAGTAAGAG TAAGCATTGT
8101 TTACTAAACT TTGGTTTTAT TTAAGTACAT ATCTATATAC TATGTCAGTG
8151 AGAAACATTT CTCCACTTCA TGTTTGAAAA ACATTCAAA AGCTAAGAAA
8201 AAGTTTGAAA ACCTGTTTGT AAGTACACCT GGGGTAAAGG TACACCCTGT
8251 GGCATAAGAT GTCGGGAACA ACTGAGGGTA AGAATGGGGA TGCATTACTA
8301 TCGTAAACTT CTGCTAAAGC ATAAGGATGT GAGTGCTGGG AGCAAAGCAG
8351 TGCTCACCAC TTCTGCAATT TTCTATTGCA GCATTTTAAA TAATATGGGA
8401 AAAAGTGGAC TGCAACCAAA GGCAAAGAGG GATGGTGATG GTGAAGGGTA
8451 AGATTGTATT TATTGTCCAA AGGCTAAGTG CATATACATA TGTGTTTGGG
8501 AGAAGGCATC ACGTAATAGT TCTTAACCTA CTCTGAGAGA AGGTTGTCCA
8551 CATTTCTTAA AGTATACATG TAAACCAACA ATGAAATTAT TTTAGTGACT
8601 TGAGAATCAA AGTGCTAGAG TTTGAATCCC TGTTCTACTA CTTGCTAGCG
8651 GTGTGACCTT GGGCCTGTTT AACTCTTGAC ACCTTGTTTT CCAAATTTAT
8701 AAAGTGGAGA TAATAATATC TGTCACATTG TGTTGTTGTG AGGATTATAT
8751 GAACTAATAT ATGTAATGTC CTGAGAACAA TGTCTGGTAC ACATTAAGTT
8801 AATTAAAATT AGCTGTTCTT ACTGTTATTA TTAGACATGA GCTAGATAAC
8851 AGTGGCCTCT ACATGTGAAA GATTATTTTA ATTCTGATGT AGTTCAGTTT
8901 ATCTATTTTT TTTATTTTTG TCCCTTTTGC ATTGATGTCA TATCTAAAAA
8951 ACCTGCCTAA CTCAGGATCA CAAAAATTTA CTCCTGTATT TTATAATTTT
9001 AGCTCTTTAG ATCTAGGATC CATTTTTAGC TAATTTTTAT ATATGGTGTG
9051 AGGTAGGGGT ACGTTTCAT TCTTTTGCAC GTGAATAGCC AGTTGTCCCA
9101 GCATCATTTA TTCAAAGAC TATTCTTTCC TCACTAGAAA AAATATTTCT
9151 TTAAAGAATA ATGAATCCTT TTTTTTTTCT TTTTAACCGC TGTTACTCAG
9201 TTGGAAAAAG AATAATGAAT AATTTTAAGT AATTTTCCTA CAGGTAAATT
9251 TAAGTCTTTA TGTTTAGATT ACACATATTA GGAAATAATG GATTTGTATT
9301 CCATAGGTAT GCTTGATCTT TATAAAGTTC CCTGTCTCTG GAAAAACTAA
9351 AATAAGGCAA AACAATCTTC TTAGTAGAGT TATTTTTACA AGAAAGTTGC
9401 AAGCCAGTTT TAGTTCATCG ATTGGATAAT TTTTCCTGCT TGCTGGAGGT
```

<div align="center">FIGURE 3, page 3 of 27</div>

```
 9451 ATTTCAGTAT TGGTAATACC TGAACTATGA GGATGCATGA ATGATGCATT
 9501 TTAGGAATTT GTTTCTGTGT CCATACCAGG CATAATGAAT TAAGTTATCT
 9551 GTTAAAAATA CAGGATTTTT GCTCAATATA CAGTTGTAGA AGAACTCATT
 9601 GTCCAAATTT TTAAGACTTT TTTTTCTTTT TTTTTTTGAG ATGGATCTCG
 9651 CTCTGTCGCC CAGGTTGGAG TGCAGTGGCA CAACCTCCAC TCACTGCAAC
 9701 CTCCACCTCC AGGGTTCAAG TGATTCTGCT GCCTCAGCTT CCCGAGTAGC
 9751 TGGGGACTAC AGGCATATGC CACTATGCCC GCCTGATTTT TTTTAGTAGA
 9801 GATGGGGTTT CACCATATTG GCCAGGCTGC TCTTGAACTC CTGACCTCGT
 9851 GATCCACCCG CCTCAGCCTC CCAAAGTTCT GAGATTACAG GTGTGAGCCA
 9901 CCGCGCCCGG CCAGACATTT TTTTTTTTT TTTTTTTTT GCTGTCTTTG
 9951 TCATATTGTT AGTCTTTTGG TTAAGCGATA TTATAACTTA GTCATATGAG
10001 TAATATAATG CAACATGCTG AATTGTGTGT GTGAGAGGGG GTTGTTTTTT
10051 GTTTGTTATT TGTTTTTTAA ATAGAGATGA GATCTCACTG TGTTTCCCAG
10101 GCTCCCTTGA ACTCCTGGGC TCAGATGATA TAGCCTCCTG CCACAGCGTC
10151 CTGATTAGCT GGGACTACAG GTGTGCACCA CTACACGTGG CTTTCCTGAT
10201 GAAATTTTAA ATACCCAAAT ATTTGAGCAG AAATAATAGC TTGTGTTTAT
10251 TGTTTTTCTA CTATCTGTCA AGTATAGTAT TAAATGTTTT ACATAATTTG
10301 TCTCCAGTCC ACATACAATA CTCTAGTAGA AGTGGGTAAC AAAACCAAGG
10351 TACTCAAAGA GGTTAATAAG TAACTTGCGC TGGATCACAG AACTAACGGG
10401 AGGCAGGGCT GGAATTTGAC TCTAGGTCTT TCTGACCTCA AAGTGCAGTA
10451 AAGTCATGGA ATTTCTCTAC TAGGCCACCT GGAAGAAAAG TGATCTTTTT
10501 TCCAGTCTTT TTTGTTACTG TTTTTCAGCC AGGAGATAGT AGAGTTAGGT
10551 AGTAGAATAG TAGTCACTGG CATCCGGTAG TCAGCCCTCC AAAAAAGTTT
10601 TTGATTTTTT TTTTTTTTTT TGTCTTAAAC TTGGAAGCTA CTAACTTTCA
10651 GGTCATACTT TCTTATCATC CAAGAGCTGG ATATTTAGGT AGCAGAAACT
10701 ATGGAATTAT CCTAAGTCCT CTTGAAGCTT CAGCTGTTAA AATTAATTGG
10751 TTCTGATTAA CACTGTGCTC AAGATTTACA TTTCTAGGAG CCACAGTTTG
10801 ATTGGTCTAA CTTGGATCTA TGTGTTTTCT TTAGCTGGGG AGGAGAAGGT
10851 ATCTTGATTG ATACCTTCAC CAGGACTGCA TGCAGTGAGG GACAGAAGTT
10901 TCCTTAAAAT AATTGGGTTC TGTTATAGGA AGAAGGGGAA GGAGATACCA
10951 AGTGGGCAAA ACAATCAGGT TCTATTACAT AAATAATAAA CCTAATGTGA
11001 CGATAATAAA TGGATAATAT GATTATTTTA AGTTTGGAAA TATACCTGGT
11051 TATTAGTATT GGATATCTGG TAGTGGGGTT GGAGAAAAAG TCGAGAATAA
11101 GAAAAGACTT AAAATCGTAA AAATTAACTG GAAAGAGGA TGGCTGAGCA
11151 GATACATATA TGTTAGATAA TGTTCATAAT GGCAAACCAA CCTGAAGATT
11201 TGTTTAAATT GTAGTATGTA GCCAGGTGTG GTGGTGCTTG CCTGCAGTCC
11251 CAACTACTTG GGAGGCTGAG GCAGGATGAT TGCTTGAGCC TAGGTTTGAG
11301 GCTACAGTGA GCTATGTTTC CACCACTGCT TTCCAGCCTA GGTGGCAGAG
11351 CAAGACCCCA TCTCTAAAAA AATAAAGTAA AATGAATAAA TTATAATATG
11401 TTATGACAAA TATAGTTATC TGAAGTCACA GAAAATGTGC ATGTGCATTT
11451 AATGATGTGA AATAATTTTT AGGAAGTATG AATAAAAAAA TCAACTTTTA
11501 AGTGTGGCTA GTATGATCTT ACCTGTATCT CACTTATAGA AAATATAAAA
11551 GGCTGAAGCC AGTCACCAGT TTAATAGTTC TAACCTCTTG TTTACTTGAT
11601 TCCCTTTTTT CTCCTCCCCA GCAATCCTCA TATAGTTAGG TAAAGTTGGT
11651 TCTTCATCAG GCTTGTTGCA GAAACCCCTA AGCCTTTTTA CTTAAAGCTT
11701 TTTGAAACCC AGAAACCCAT CTTTTGAATT CAAAAGTTTT GACTGTTATT
11751 AGTCTTTTTG TATGTTTGTT GGCCGCATAA ATGTCTCCTT TTTATGAACA
11801 GAGAAGTGTC TGTTAATATA CTTTGCCCAC TTTTTGATGG GGTTGTTTGT
11851 TTTTTTCTTG TACATTTGTT TAAGTTCCTT GTAGATTCTG GATATTAGAC
11901 CTATGTCAGA TGGATAGATT GCAAAGTTT TCTCCCATTC TGTAGGTTGC
11951 TTGTTCATTC TGATGATAGT TTCTTTTACT GTGCAGAAGC TCTTTAGTTT
12001 AATTAGATCC TATTTGTCTG TTTTGGCTTT TGTCGCCATT GCTTTTGGTG
12051 TTTCAGTCAT GAAGTCTTTG CCAGTGCCTA TGTCCTGAAT GGTATTGCCT
12101 AGGTTTTCAT GGTTTTGGGT TTTACATTTA AGCCTCAAAT CGATCTTGAG
12151 TTAATTTTTG TATAAGGTGT AAGGAAGGGG TCCAGTTCCA GTTTTCTGCA
12201 TATGGATAGC CAGTTTTCCC AGCACCATTT ATTAATATTA AATAGGGAAT
12251 CCTTTCCCCA TTACTTGTTT TTGTCAAGTT TGCTGAAGAT CAGATGATTG
12301 TAGATGTGTG GTGTTATTTC TGAGGTCTTT GTTCTGTTCC GTTGGTCTGT
12351 ATATGTGTTT TGGTACCAGT ACTATGTCGT TTTGGTTACT GAGCCTTGTA
12401 GTATAGTTTG AAGTCAGGTA GTATGATGCC TCCAGCTTTG TTATTTTTGC
12451 TTAGGATTGT CTTGGCCATA CGGGCTCTTT TTTGGTTCCA TATGAAATTT
12501 AAAGTAGGTT TTTCTAATTT TGTGAGGAAA GTCAATGGTA GCTTGATGGG
12551 AATAGCGTTG AATCTATAAA TTACTTCGGG CAGTATGGCC ATTTTCATGA
```

FIGURE 3, page 4 of 27

```
12601 TATTGATTCT TCCTATCCAT GAGCATGGAA TGTTTTTCCA TTTGTTTGTG
12651 TCGTTTCTTA TTTCCTTGGG CAGTGGTTTG TAGTTCTCCT TGAACAGGTC
12701 CTTCACGTCT CTTTTAAGTT GTACTCATCA TCACTGATCA TTAGAGAAAT
12751 GAAAATCAAA ACCACAATGA GATGTCATCT CATGCCAGTC AAATGGTGAT
12801 TATTATAAAA AGTCAAAAAA GAATAGATGT GGGTAAGGCT GTGGAGAAAT
12851 AGGAATGCTT TTACACTGTT GGTGGGAGTG TAAATTAGTT CAACCATTGT
12901 GGAAGACAGT ATGGCGATTC CTCAAGGATC TAGAACCAGA AATACCATTT
12951 GACCCAGCAG TCCCATTACT GGGTGTATAC CCAAAGGATT ATAAATCATT
13001 CTGCTATAAA GACACATGCA CACGTATGTT TATTATAGCA CTATTTACAA
13051 TAGCAAAGAC TTGAAACCAA CCCAAAAAGC CATCAATGGT AGACTGGATA
13101 AAGAAAATGT GGCACATATA TACCATGGAA TACTATNNNN NNNNNNNNNN
13151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14501 NNNNTAAAAG ATACATCCTT TATTCATGCG TAAGATGAAA TCGAGAGGTG
14551 AAATTGGATA TACTGTTGCT TTTAAAAAAT TTTAACATAT ATGTAATTTT
14601 TTGTACTTAT CTCATTTTAG CCTATATAAG TTATATATAT TTTGTTTGTT
14651 TGTTTGTTTG TTTTGTTTGA GATGGAGTCT TGCTCTGTCA CCCAGGCTAG
14701 AGTGCAGTGG TGCAATCTCG GCTCACTGCA ACCTTCGCCT CCTGCATTCA
14751 AGCGATTCTC CTGCCTCAGC CTCCTGAATA GCTGGGATTA CAGGCACCTG
14801 CCACCGCGCC CAGCTAATTT TTTTTATTTT TAGTAGAGAC AGGGTTTCAC
14851 CATCTTGGCC AGGCTGGTCT TGAACTCCTG ACCTTGTGAT CCATGTGCCT
14901 TAGCCTCCCA AAGTGCTGGG ATTACAAGCG GGAGCCACCG CGCCCGGCTG
14951 TAAGTTATAT CTTACACAAA TCTAGGTTTC ATTCAGAGAA TTATATGCAA
15001 AGAAACAGTG CAATAGGATT ATTTTAAAGC TATTGTTATT GTTAGAAAAC
15051 ATAATACCTT TAAAATTCCT TTTCACATTA GAAATATAGT GGCTTCTCCC
15101 CAGTTTAGGA TAGAAATTTT CCTTTTCTTC TCCTTCTTTA TACTATTCAG
15151 ATTTGCATGT TTGACAGAAC AAATTATAAG AGAAAATATT TGAAATGTCA
15201 CATACTAAAG TAAATGTTTG AATGTTTGAA AATTTCTGG TTTTCAGAGA
15251 TTTTGAATTG CTGAATCGTT GTGTAAATTA AGATGTTGAG TAGTTTCCAC
15301 AGAGTAATTA TTTGAAAGTC ACTGAAAGCA AGACACATGC CTAATGTAAA
15351 TGTTTATTGC ACTACTGTAC CTTTTTCTAC CTCATAAAAA TGAGAATAGC
15401 AGTCTGTACT TTTCCACTTC GTCATTCGTA AGTCTTTGCA GAAATTCATA
15451 TTTTGTTTGC TTATTATCTT CACGCTGTAA ATAGCTTGAA AATTCTTTAA
15501 GTGGGGCTAG CGATGTATTA TGGATACATG TTAAGTGGTA TAGAAATTTC
15551 ACTTTTTTTT TTTTGCATAA AGAGTAACAA GACCAGTAGT CCATATTTCT
15601 TCAGCTCTAC CCAGAGAAGG GCAATGTAGG AGGGAAAATG AAGTTTGCAA
15651 AATATTTCAT AGTAGGCTTT TTCTTAAAGT AACTTCAGAC TTACAGAAGT
15701 TTAAAAATAG TACAAAGAAT CCCCATATAC CTGTCACCCC AATTCCTGAA
```

FIGURE 3, page 5 of 27

```
15751 ATATTAATAT TTTACCACAT TTGTTCATTA TGTCTGTATT CTCCAAGTAC
15801 GATATATGCC ATTATATGTA ATATGTAGCA TTTTATATAG ACATAGGGCA
15851 TGTATGCACT ATATATTTTT TTCTGAGCCA CATAAAGAGT AAAACGCAGA
15901 CATGACGTGC TTTTACTCCT AAATACTTCA GTGTGTGTAT TCCCTCAAGA
15951 AAGGGCATTT TCTTCTGTAT AGCTACCGTA CACTTCTACA CTTTTCAAAA
16001 TCAGAACATT TACATTGATA CCATACTATG ACATGATCTG CAGACCATTT
16051 TCCAATATGC CAGTTGTCCC ACTGTGTCCT TTAGTACAAA AGAAAAAAGT
16101 TTTTTTTCCT GGTCTAGGAG CTAATCCTGG AGCACATGTT ACATCCTGTT
16151 GTTTTAATCT AGAACCGTTC CTCAGTTCTT TATCTTTCAT AACCTTGACA
16201 TTTTTGGAGA GTACAATCCA TATATTTTGC AGAATTTCCC TTAGTTTGGG
16251 TGTGTCTGGT TTTTCCTTAT AAGATTCATT TTATGCATTT CTGGCCAGAG
16301 TACCACAGAA GTACTGTATA TCTTACCAGA AAGCCTAAGT GGCATTTGCA
16351 TTTTCTAAAT GATCAATTTT AATATTATAT GGAAAGCAGA GTCAGAGATT
16401 CTCACATATG TCAAGATATT ATAAGTATTC CTGTTATATT TATTCTCCAA
16451 TTGCTTTTTC TCAAGAAAAT TTGTGGCCTT TCAGCTAGCT TTTCAAAGTG
16501 GAAGTTACTA CATAACATTA GGATGGGAGG GGTGGGGAAG AGCTTTATTA
16551 AAGCTTTAAG ATTGAGCTTT TGAGTATGTG TTGTATGTAA ATGAAAGTGG
16601 GCATTGATGC AGGGATTGGG CCTTTAAACC TTTGGCCAAG AATGGTATCA
16651 ATTATTATTA TTATTATTTT TTGGAGTACT TCTGCTAAAA CACTGAAATC
16701 AGTGTGCCAC TCTCCTTTTA GAAGTTTTAC ACCTTTCCAA GGTACACTTT
16751 TTTTTTTGGA GACGAGTTTT GCTCTGTCGC CCAGGCTGGA GTGCATTGGC
16801 GCAATCACAG CCCACTTCAG CCTCTGTTTC CCAGACTCCA GCAGTCCTTC
16851 CACTTCAGCC TCCCGAGTAG CTGGGATTAC AGGTGCACAC CACCATGCCC
16901 AGCTAGTTTT TGTAGAGATG GGGTTTTGCC CATGTTGCCC AGGCTGGTCT
16951 CCAACTCCTG CGCTCAATCT ATCCGTCCTC CTCAGCCTGC CAAAGTACTG
17001 GGATTACAGG CGTGGGCCAC CACTCCCGGC TTCCAAGGCA GGCATTTAAA
17051 TGTAATAAAT AGGGAGATAA GCAAGAACCC TGTTGGACCT GGTAGAAGCA
17101 AACATTTATT AGTACTATTA CGTTGTTTAA AATATTAGCG CCTTCTATAT
17151 TCATGTCCTC CCAGAATTAT CAAAAAACCT ACTCTATAGT TTATTTGGCT
17201 TATATCTCAG GAGTAATAAA ATTAGTTAAT AGTATTGGCA TCGTGGTTCT
17251 TTGTGTATTC CTCCCTTATC CCACCCCAAG TTGATTTCAC ATGATCTCTT
17301 GATCTAGTCT AAGAATGTTT ATAGTGATTA CGAGAAGTTC AGATTCTGGC
17351 TTTAACATAT ATAATTGTTT TTTAATCTGT AAACCAAAGA GAATGAGTTT
17401 GTTTAAACTA GAAAGATGGC AAGAGTAGTC TGGGAATTTT GTTCCATTCC
17451 TTAAAAGTCC TATAATAAAA TAAACATATC TTGTGTTTTA TTTTTACAAT
17501 TTTTTTAAAC ATTAGTACAG AGTGCCACTT CTTATATTCT ATATCAAATA
17551 ATGAGCTACA TTTTCAATAA TAACCTCTGA GTAATTTTTG GCATTAAAAT
17601 GCTGCATTAC AAAATAATTT GAGGATATAA TTTATAATCA CTTATGCTAA
17651 AATCACCTAT TTGAAATTAT GTATGAGGTT TTCAAAGTTT ATAGTGCTTT
17701 GGAAAAAATT TAAATGTTTC TTTGTTTATG TATCTTTATT ATAAGCTGTA
17751 GCATATATCA TGTAGTTGTC AAGGATGCTG ATAGATACTT AATATTTAAA
17801 GGAGACTTGT CTAAAGTTAG CTGTCCAGGA CTAGAATCTG GGCCTTTTGG
17851 TAACAGCTCA TTGCTCTATT TACTTAAATG ATGATTGGAT TCGTTAGAAT
17901 TTCTCTATTT TCATAGCTGT CTCTATGGTT CTATGAAAAT ACTGTGTGTG
17951 TGCTTATACA TATATGTATA CCTGTAAGTA CAAAGTAGAA AATGAAAGTT
18001 CATTTTCTGC TTTTGACAAT TGTAATCCCC AGAGATAACC GTTATTAATA
18051 TGTTGTCTCA TGTTTGGTCA TACTGTTTTC TCTGTATTCT GTGTATTACT
18101 GTATAAATTT TACACAGTAA TTTGCATATT AAAAATGCTG GTCTACACCT
18151 GGCCCTTTTT TAAAAACTGC AATTATTAT GGCCAATTTT TTATACCAGT
18201 ATATATTGAT CAACCTTATT CTTTTTAACT GCTGCATTTC ATTCATTACC
18251 AATAGATGAG ACATTTCCAT TGGTTTGAAT TTTTCAGTAT TACAGATAAT
18301 GGTTCAATTA AATATTTAAG CTTTTGTGCA CTTGTAGAAT TAATTCCTAG
18351 ACATAGAACC CTTATATTTT GATAGGTATT TCCAAATTTC TTCCCAAAAT
18401 GTTTGTATCT CTTTACTTCC ACTCTCAGGT CTAATAATTT TCACTTGGAT
18451 TATCATATTT CTTACCCAGC CTGTTTTTTA CACTCTAAAC TCTTTTTCTT
18501 TTCTTTTTTT TTTTGAGACA GCATCTTGCT CTTGGCCCGG TTGAAATGCA
18551 GTGGCACGAC GACCAACCTG GGCTCAAGCA ATTCTCTCAA CTTAGCCTAC
18601 TGAGTAGCTG GGACTACAGA CACATCATCAC CATGCCCAGC ATTTTTTTTT
18651 TTTTTTTTTT GGATTTTTAG TAGAGATGAG GTTTTGCCAT GTTGCCCAAG
18701 CTGGTCTCAA ATTCCTGAGC TCAAGCAATC CACCCATCTC AGCCTCCCAA
18751 AATGCTGGGA TTACAAGCGT GAGCCACTGC ACCTGGCCCA AAAGCTCTTT
18801 TTCTAATAGC AATATAAATT GTCTTTTACA GACTATACTC ATATATGTTT
18851 CTTCTTTCAG AAATAGGTGT TAAGTGTATC TAACATGGAA TGTATAGCTA
```

```
18901 TAATTCTCAT TGTGAAACCA TAGCCTAATT TATTTCATAT TACAATTTAA
18951 AATTCATATT TTTTAGGAAG TTTTCTTAGA TTAATCCGCC TAGTTCCAGG
19001 TGCTACAGTC CCAAGATTTC TTTCTTTTTA ACAAATTAAA TATAGGTAAC
19051 ATGACTAGAA TTGTAGTCAA AGAATATTGG AACCTTGGAA CTTCAGTATT
19101 TGAACTTTAT TTTGAAATAT AATTTGTTAT ATTATAAAAA TATTATAATA
19151 TATTGCACCT GGAAGTTAGG GGCAGTTTTT TTTAATTCTC TTTGTATCTG
19201 CTACACTGTA AAGTGCTATT TATGTAAAAA ATTCTTAATA GAAGTCTTCA
19251 GTTGTAAAGT CTGCTGTACA GACTTTAGAT CAGGGATTGG CAAACTATGA
19301 GCCATGTGCC AAATCCTGCC CTTCACCTGT TTTGTAAATA AAGTTTTATC
19351 AGAACACATT CAGACTCATT CATGAACATA TTGTCTATGA TTTATTTTCT
19401 GCTACTATGG CAGAATTGAG TTGTTGCAAC TGTGTGGCAT CCAAAGCCTA
19451 AAATATTTAC TCTCCTGGCT CTTTGCCAAC CCGTTTTAGA TTATGAGCAC
19501 TTTGGCATTA TTATGTTTTT GTTTTCTTTC TATAGCACAC AGTAAGATGT
19551 TCTGCCCACA TTGTGCATAA TTTATGGGTT TATTCAAGGA TTTATGCAAG
19601 TGTAGCTGCA AGAAAAAAAC CTAGAAGTGA ACTTGCTAGG TTGAAGAGCA
19651 TCTGTGTATG TTAAATTTTG TTAGCTTTCG CCTTCCCAAA GGGATTATTC
19701 CATTTCATAC TTAAACTACT AATTTTGTGA TAGGACTTCT TTCTCCATAG
19751 CTTTGCTAAA TTAATGCATT CACACACTTC ATCTTTACTA ATCTGATAGA
19801 GGGAAATGAT ATTGTGGATT TGATTTGCAT TTCTTTTTAT GTGTTAGCTT
19851 GAGCTTATTT TCATATTTAA AAGCCAATTG TATTTCTTTT TCTTGAGCTA
19901 TCTTTTAATG TCCTTCCTGA TACATTTCTG AAGTCTGTGA TACTCATATA
19951 AGATATATGG TGAACATGTG TCAAAGATTT ATTTGACTCT AATGAGGGAA
20001 CCCGCCTGAT GACAAGGCTG ATTGAGAAGA GGATGTGTGA GATGAAGTGT
20051 ATATCATCAG TGAAAGAAAG CAAATTCTTA CAGGGCAAAA ACAAAACCAC
20101 AACTCTAAGG GTTATTGTTT CTACTGGACA GAATTCATTT GCATTTTACC
20151 AGATAAAAAT TACTATTTTC AATTTATCTT TTACAAATCA TTTTCTAATT
20201 TTACAGAGTC TATTCCCTAA TCAGTAGTAA ATAGTCTTCA AAATTCTCCG
20251 CAGCGTCAGG TGACTATTAT GCAGGCTAAT TGTTGACACT CGGGCTTGAC
20301 TTTAAGAGAA CATGCCATAA TCTTTTGGCC TTACTTCCAA GTTTTGGATA
20351 ATTTTTCTTA ACACATTTTT CTCTAATTGC AATGATTTCA AGTGATATTA
20401 TTTCTTTTTT TTAAATTTTT TTACTATTTA TTGATCACTC TTGGGTGTTT
20451 CTCGGAGAGG GGGATTTGGC AGGGTCATAG GACAATAGTG GAGGGAAGGT
20501 CAGCAGATAA ACATGTGAAC AAAGGTCTCT GGTTTTCCTA GGCAGAGGAC
20551 CCTGCGGCCT TCCACAGTGT TTGTGTCCCT GGGTACTTGA GATTAGGGAG
20601 TGGTGATGAC TCTTAATGAG CATGCTGCCT TCAAGCATCT GTTTAACAAA
20651 GCACATCTTG CACCGCCCTT AATCCCTTTA ACCCTGAGTT GACATAGCAC
20701 ATGTTTCAGA GAGCAGGGGG TTGGGGGTAA GGTTATGGAT TAACAGCATC
20751 CCAAGGCAGA AGAATTTTTC TTAGTACAGA ACAAAATGGA GTCTCCTGTG
20801 TCTACTTCTT TCTACACAGA CACAGTAACA ATCTGATCTC TCTTTTCCCC
20851 ATATTTCCCC TTTTCTATTT GACAAAACTG CCATCCTCAC CATGGCCCGT
20901 TCTCAATGAG CTGTTGGGTA CACCTCCCAG ACAGGGTGGC GGCCAGGCAG
20951 AGGGGCTCCT CACTTCCCAC ACTGGGCGGC CGGGCGGAGG CGCCCCCCAC
21001 CTCCCAGACG GGGCGGCTGC CGGGCGGGGG CGCCCCCCAC CTCCCAGACT
21051 GGGTGGCCGG GCGGAGACGC TCCTCACTTC CCAGATGGGG CGGCTGCCGG
21101 GCGGAGGGGC TCCTCACTTC TCAGATGGGG TCGCGGCTGG GCAGAGGTGC
21151 TCCTCACCTC CCAGACAGGG TGGCGGCTGG GCAGAGACGC TCCTCACCTC
21201 CCAGACGGGG CAGCCGGGCA GAGGCGCTCC TCACATCCCA GAGGGGGCGG
21251 CCGGGCAGAG GCGCTCCCCA CGTCCCAGAC GATGGGCGGC CGGGCAGAGA
21301 CGCTCCTCAC TTCCTAGACG GGATGGCGGC GGGGAAGAGG CGCTCCTCAC
21351 TTCCTAGATG GGATGGCGGC CGGGAAGAGG TGCTCCTCAC TTCCTAGACT
21401 GGGCGGCCGG GCAGAGGGGC TTCTCACATC CCAGACGATG GGCAGTCAGG
21451 CAGAGACGCT CCTCACTTCC TAGTACAGGG TGGCGGCCGG GCAGAGGCTG
21501 CAATCTCAGC ACTTCGGGAG GCCAAGGCAG GTGGCTGGGA GGTGGGGGTT
21551 GTAGCGAGCC GAGATCACGC CACTGCACTC CAGCCTGGGC AACATTGAGC
21601 ACTGAGTGAG CGAGACTCCG TCTGCAATCC CGGCACCTCG GGAGGCCGAG
21651 GCGGGCAGAT CACTCGAGGT CAGGAGCTGG AGACCAGCCC GGCCAACATG
21701 GCGAAACCCC GTCTCCACCA AAAAACACAA AAACCAGTCA GGCGTGGCGG
21751 CGCGTGCCTG CAATCCCAGG CACTCGGCAG GCTGAGGCAG GAGAATCAGG
21801 CAGGAAGGTT GCAGTGAGCC GAGATCGCGG CAGTACAGTC CAGCCTCGGC
21851 AACAGAGGGA GACCGTGGAA AGTGGGAGAC GGAGACGAGG GAGAGGGGGA
21901 GACCGTGGAA AGCGGGAGGT GGAGACGAGG GAGAGGGAGA GGGATTATTT
21951 CTGTATGACT TAATAATGAA TTTCTAAGAG GTCACTTAGC TCACTGTTGT
22001 CTCTTCTAAA ACATACTCAT CTTTCCTTTT CTCTTCTGTA GGAACTCATT
```

FIGURE 3, page 7 of 27

```
22051 ATACAATGAC AAATGGAGGC AGCATTAACA GTTCTACACA TTTACTGGAT
22101 CTTTTGGATG AACCAATTCC AGGTGTTGGT ACATATGATG ATTTCCATAC
22151 TATTGATTGG GTGCGAGAAA AATGTAAAGA CAGAGAAAGG CATAGACGGG
22201 TAAGTGTTTT TAGTAAAAAT TTTTAAAAAC ATAGTGCATA ATTAGATCTT
22251 TTAATAATAT ATTTCTGCCA ATGATCTCAG GCTGCCAAAT GTTTACATTT
22301 AATATAAGTA AATGTCTACA TTTCATATGT GGTACATGTT TTTTTCTTTT
22351 TCTATGTTTA ATTTTTTTAG TTTACTTATA CCCTGTAACT TTCCAGAAAG
22401 GATTTCAGGT AGCTAAAAAA CAAAGAAATA CAATAAGAAG ACAAAATAAG
22451 AAGGAAAGGG AAAAATACAG CACAGGAGTT GGGGGGAAGA ACAAGCCAAG
22501 TTCCAGATAT GGAGGTCAGC ATGATTTTGG GCTTTGAGCA GCCCACCAGC
22551 TAAGGCAAAA AAGGAAACTC ATTGCATAGC TCTTACCTAT GGAAAAAGAA
22601 GAAATCTACT GGGGGCAGAT GGTCTTGTGG GATTTTGCTG TTTTCTTTTA
22651 TCTCCTTTCC CAGCATTTGA TTCTGAGATA TTTCTCAATT TGGCTCCCAA
22701 ATAAAGCTTA TTGAGTGTTG TAATGGTTTA CTGTTTTTTT TAAAAATGGC
22751 TTTAACATAT AAAAGTACAA CTTATGGATC CTTTTTGTTT GTGGTCGTGA
22801 CTTACTGATA ATATAATCCA AAATACATTT TTTATTTTGT ATTTATTTAT
22851 TTATTTTTGA GACGGAGTCT CAGTCTTCTG CCCATGCTGG AGTATAGTGG
22901 TGTGATATTG GCTCACTGCA CCCTCCGCCT CCTGGATTCA AGCGATGCTC
22951 CTGCCTCAGC CTCCTGAGTA GCTGAGACTA CAAACGTACG CCACCATGCC
23001 TGGCTAGTTT TTATACAAAA TACGTTTTTT AAAAAACAAT TTTTTTTTTG
23051 GAGGTCGGGG GACTGTCGCC CATTCTGTTG CCCAAACTGG AGTGCAGTGG
23101 TGCAATCTTG GCTCACTGCA ACCTCTGCCT CCCAGGTTCA AGCGATTCTT
23151 GTACTCAGCC TCCTGAGTAG CTGGAATTAT AGGTGTGTGC CATCATGCCA
23201 AGCTAATTTT TGTATTTTTA GTAGAGATGA AGTTTCGCCA TGTTGGCGAG
23251 GCTAGTCTCA GACTCCTGGC CTCAAGTGAT TGGCTGACCT CAGCCTCCCA
23301 AAGTAGAAAA TCTTCTTGAA AAATAAAATT CCAAATCTCA AAAGGCCCTA
23351 TATAATTTTG GTGTTGGAAA TTTACTTGTC AATGAAAATG ACTATTTACA
23401 CAAATTATAA GCTTCCATAT TAATATATAT GTGTGTGAAC CTGAAATTCA
23451 AATTTTATTA TATTGTTTAT GAAAGGTACA GCCTCTGAGA TTCATCAGAT
23501 GGTATTTACC TTTAGGGCAT ATCTAAAAAT AAAATACAGT ACATGAAATC
23551 CAGTGCTTTA ATCCAGTGAT TCTTAAACTT TTTGCTCTCA GATCCCCTTT
23601 AAACTCTTAA AAGATATTGA AGAGCTCCAA GGAGGCTTTG TTTACGTGGT
23651 TTTTATCAAT GGATATTTAC CATATTAGAC ACTGAAACTG AGGATTTTAA
23701 AAAAAAATAA TTCATTTAAA AATAACAGTA ACAAACCCCA TTACATGTTG
23751 ACATAAATAA CATTTTTACG AAACTATATT TTCAAAAATT AGTGAGAGAA
23801 TGACATTGTG CTACATTTGT TATAAATCTC ATTATTGTCT GGCTTAATAA
23851 AACACTGCTG GATTCTCATA TCTGCTTTTG NNNNNNNNNN NNNNNNNNNN
23901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
23951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
24851 NNNNNNNNNN NNNNNNNAAA TATTGATTCA CTGATTTATG TGGATCTTTT
24901 AAATGTTGAC ACTTATATAA TATAATACAA TATTTTAAAA ATCACATTTG
24951 TTAATTTTAC CTTTGATCTA TTCAGAAAAG ACTCTAAGTA TTGGGAACCT
25001 ATCATCCTCA CAGTGATAGA TACAAGTTTC CTAAAATTCT GATTTTTACT
25051 GGAGAGCTCA AATTCTATCA TTGGAAACAA ATACACATTT ATTTAACTTA
25101 AAAATGACAG GATTACTTGG TTTCATTATT GAGAAAATAC CTGTCAAATT
25151 CCCAAGTCTG GAAAACCATG GTTTGATGTC ACTCTTTCAA GTAAAAATGG
```

```
25201 CATTCCATGT AAGAAGTGTC TAGTTTATTA TGCAACTCAA ATAAATTACG
25251 CAAGTGCTTT TCTTTAGGAC ATAACTTCAT ACATACTTCC ACAAGCAGCA
25301 GATGTGTGTA GTTATGCATA GTTCCTTATG CATGGTTCTT ATTTCATCAC
25351 ACAAAATATT AAAAAGACTC AGTGATTGAG ACGTAGCAGT TTTTACTGCT
25401 TCATCAAAGA TGCTCTTATT TGAAACTGGC ATAATATGAT TTATTTATTT
25451 GATTTTACTG GGAAGCATGG CAGTCAAGAA TGTAATGACT GCCAGTACAT
25501 TTGAGTGCCA CTGCTTGATT TTTGCTATGG AGTCAGCAAT TTTGCCACTG
25551 GTTTTGCATT TTCAGTAAAA ATGTCAACAC AGTGAAAAAG GCACATAATG
25601 TCTTGTATTA TTTTGTAAAC AGTTTTATCT TGCAGACCCC TTGAAAAGGT
25651 CTCGGGGATC CTCCAAGGTG CCAGTAGACC GTACTTTGAA AATCACTATT
25701 TTAATCCAAA GTGCCTAGAT CAGACACACT ATAAATCCTG TGTCTTGTAT
25751 GATCATTAGG TAAATACATT TGTACTTAGA AGTATACATT CAGAGACATT
25801 AACAGTATTC AGGTTGGGAT TTAAGTATAT TTTAAAGTGT GGTACCTAGA
25851 GAGTATCCAT GACACTATGT TCATAAAATT TTAGAGAAAA CTGAGATCAA
25901 AGGAAACCAA AACAGGCTGG TCATAGTGGC TCATGCCTGT AATCCCAGTG
25951 CTTTGGAAGG TTGAGGCAGA GGATCGCTGG ATCCCAGGAG TTTGAGACCA
26001 GCCTGGGCAA ATATGGAGAC TATCTCTACA CAACAAAACA AAAATTAGCT
26051 GGGTATAGTG TCTTGCGCCT ATAGTCCTAG CTACTCGGAA AGCTGAGGTG
26101 GGAGGATCCC TTGAGCCTGG AAGTTCTAAG TTACAGTGAA TTATGATTGC
26151 ACCACTGCCC TCCAACCTGG GTGAAACAGC AAGACCCTGT CACCCTCCAA
26201 AACAAACAAA AAACACTTTT TTCTCTGAGT ATGTAAATGG TTAGTGTACA
26251 GTCCTTGAAA ACATTGCAAA TAGTATAGCA ATATATGAAG TAGCCAGTAT
26301 GTGTCCTAGC TAATTTTATC AATCATCTCT TCCTAGACCA ATCAAATATT
26351 TTTCAATATT TTGATCCATG CTTATATGAA CAAGATTTTT TAAAGCTGGA
26401 AAATTCCACA CATTTATATA CTTACTATTG TTCTTAAAAT TAATTTTTTT
26451 TTTTTTTTTT TAAGCAGAGT CTTGCTCTTT TGCCCAGGCT GAAGTTCAGT
26501 GGGGCGATCT CGACTCCCTG CAACCTCTGC CTTCCAGGCT CAAGCAGCTC
26551 TCGTGCTTCA GCACCCCAAG TAACTGGGAT TACAGGCATA CGCCACCACA
26601 CTGGCTAATT TTTGTAGTTT AAGTAGAGAT GTGGTTTCGC CATGTTGGCC
26651 AGGCTGGTCT CAAACTCCCG GCCTCAAGTG ATCCACCTGT CTCAGCCTCC
26701 CAAAATGTTG GGATTACAGG TGGGAGCCAC TGCGCCCGGC CTACATTAAA
26751 TTTTAAAGCC TTTCTATGTC AGTGCATATA CCCAACCTAA TTCTTTTTTT
26801 CCGTGAACTT TTTTGTTATG CTTGTAGCCT TCCTACCCCA GATTATTTCG
26851 AAGCAAATTG TCATTCTGTA ATTTCAAATA TTACTATTTC AGTATTTTAC
26901 AAAATGGTTG CAGTTTAATT GTTGTTCCTT TTTTATTTAT TAGCTTGCAT
26951 ATTTCTATAG AGAGTTTACC CCACATCAAC CATTTGGATT ACCTGAAGTA
27001 AGGGTGGTAC AGGAAAGGGA GAAATCTTGA AATACTAGGT TCCTTAGCAT
27051 CCTCAAAGTT GACCAATGAG ATTTTTTGCT TGTTTGGTTG TTTTTTTCTG
27101 TGTCTTCTGG ACTCATGGAT TTAAGTATAT TTGTGGTTTA ATCATCACTG
27151 TTATTATTCT TATTGATGTT CATGTTATTT TAGATTAGTG GGAGCTTTTT
27201 TAGTTTGCTA TCTGTGTCCT TCGTCATGTC CTTAGATAAT CCTAATCCTA
27251 ATCCTGATTC ATCGTAGACA TTTCCCGCAG CAAACCTGGA ATCAGCCATT
27301 TCTCAAGGAG CTCTCTGATT CCATTGAAGG AAAATATAAT ATAGGTACAA
27351 TCTAGGCACT AGGTGATACT TGTTACTTCT GGGTTGGCTA TTGTTTCTAG
27401 CCTCCTAAGT TTATATGACT GTACTAATTT GAATTCATAA CTATGGGACT
27451 AAACTTCTAA TTCTTAAATC TGCATTTCCT TTAAGTCATG CCAAAAATCT
27501 GAACATCACA AACATAGTCA TTTCGTTTAC CCCACAATAC ACACATACAA
27551 CATTGTCAGT ATAACAGTAC CAACACCATC TCCAACAATA TGCCTACTGA
27601 AAAATTTTAG GTAATCTGTC TCCAGCCTCC CAGGTAGCTG GGACTGCAGG
27651 TGCACACCAC CATGCCTGGC TAATTTTTTT TTTTTTTTTT TTTTTAAGA
27701 GACTGGGTCC TTGCTATGTT ACTCAGGCTG GTCTGAAATT TCTGGCCTCT
27751 AACAGTCCTC CTGCCTTTGC CTTCCAAAGT GCAGAGATTA CAGACCTGAG
27801 CCACCACGTC TGGCCTATCC TTTATTTATT CCACCAAAGT TATTTATACA
27851 AATTACTTTG TTGTAAAGTC CCTTGGAATA GTTTCTTCTG TGGCATTATG
27901 TTACCAGTTA GATGCACCTT TGATTCATTT AACTTTACTT CAATTTTTAA
27951 GGTTTGCTTT TTAGATTTAG TTTTGTTTTA TTATACATAT ATGAAGTATT
28001 TCCACGGTTC CAAAGTTAAA TGAACAAAAC AGGCATGTTC AAAGAAGTCT
28051 AGTTTCTATC TCTGTCCCAT CCAACCCATT GTCTTCTTCC CCTTATAAGT
28101 AATAATTTAC ATTTTTAACT TGTGGTTTAT CTTCTGATTT TTAAAAATAT
28151 AAGCATAAAT ATTTATATTC CTGTCTTTTA GCATGCTTTT AGCCATCTTG
28201 CTTTTTTCCT GTATAATGCT AAATATATCT CATTCTTTTT AATTGCTGCA
28251 GAATTTCTCA TTACATAGGT ATACTGCAAT TTATTTATCT GATGCTATGT
28301 TGATGAACAT TTAAATGATT TCCAGATTTT AGGAACGGTG ATGATTGAAC
```

FIGURE 3, page 9 of 27

```
28351 TCTCTGTACA TATATCTTTT TTACTTGGTA CACTCCATCA AGCAACTACT
28401 TAAGTGACTG ACTATGATGC TGTGCAAGCA GTTATATAAA GAAAACAGCA
28451 GTGACTCAGC CTGAAAACGG CTTAATATTA TCATGTTTTC TTACACATTA
28501 TTTTTATTGA GGAAAAGCAA CATGGAGTTT AGTGATTATT TTTGAAAGAA
28551 ATAACCTATT TCTAATTCTA AAGAATGGTT ANNNNNNNNN NNNNNNNNNN
28601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
28951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
29001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
29051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
29101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNGAGTCTAG
29151 CTCTGTCACC CAGGCTGGAG TGCAGTGGCA CGATCTCTGC TCACTGCCAC
29201 CTCCGCCTCC CGGTTTCAAG TGATTCTCCT GCCTCAGCTT CCCAAGTAGC
29251 TGGGATTACA GGCGTTCGCC ACCACACCCA GCTAATTTCT GTATTTTTAG
29301 TAGAGAAGGG GTTTCACTGT GTTGGCCAGA CTGGTCTTGA ACTTCTGACC
29351 TCGTGATCCA CCTGCTTCGG ACTCCCAAAG TGCTGGGATT ACAAGCGTGA
29401 GCCACCACAC CTGGCCAAAA ATATGGGTTT CTAAAGCAAC AGTCCTAGTA
29451 CAACAGAAGA GAGGTGTTGA CTAGTTAGGG ATTTAGGTTT AGAAGTACAT
29501 TCTTAGTAAG AGAGGTGAGA CTTACCTTCT TGTGTTTTAG TATAGTGAGA
29551 TCTGGATCAA ATCTATTACT CTTATTAATC TCCTAACTTC CTACACTATA
29601 TCCAGTAGAG GACACTTTTG CCTTACACAG TAAAGAAAGA GCCTCTGGAC
29651 TCTACCAATG GGATCGGAGC TCTCCAAACC TGCATATTAA AAGGCCTATA
29701 AGTTTTGGGG GGTCCCTTTG TCCACATGAT TATTCTGTAA TACATTGTAT
29751 TTATGGACAT GGTATTATTA TACACAGATC CTGTCTTTTA AAGAACATTA
29801 TAATCCACTT AACTGCTAGG ACCAGAGAAT GACCGATAAT TCAAACCATA
29851 TTGTCTTACA GAAGACATAT ATAAAAGATG GTTATGTGTA CCAATTGAGG
29901 TTCAAATTTG ATTCAATTTA AAACAATCTA GGCCAGATTT TATATAGTTT
29951 GTGGACCCTT TGCACTCAAA TCTCAAGGTT CTTATTAAAA TGCAGATCTT
30001 GGCTGGGCAC GGTGGCTCAC ACCTGTAATC CCAGCACTTT GGGAGCCCAA
30051 GGCAGGTAGA TCATTGAGC TCAGAAGTTC AAGACCAGTC TGGCCAACAT
30101 AGCGAGGCCC AGTCTCATTG AAAGAAAAA AATTTTTAA TAAAAAATAA
30151 AAGCAGATCT TGGGTAAAGA CATGTAGTCT GGTTTACAGG TATTAACAAC
30201 TGTCTGTAAT GTAGTGATTT TGCTCCAGAC TTACCTTTTC CATTATTTAG
30251 TTCTGAAATT ACTGTTCTAT GTATGGTAAA TGAGAAAAAT TGCTAGATTC
30301 TAGAACTGTG GCTTCTATTC ATAGTTGGAA AAATGAAGCA TAAACATTTC
30351 TAATTTCAGA TCAACAGCAA AAAGAAAGAA TCAGCATGGG AAATGACAAA
30401 AAGTTTGTAT GATGCGTGGT CAGGATGGCT AGTAGTAACA CTAACAGGAT
30451 TGGCATCAGG TAAAGAAAAT TTTTCAAGCA ATCCTTTTTT AGTTAACAGA
30501 AGTATAAACT GTTCTTCCCT CCTTCCCTCA ATTTTTTTTC AGGTACCATT
30551 GGATTTTAAA AAGCATTTGT TTCTCTTCTT CAAAAAATCT CCTTAAATAT
30601 AAGACTAGGA GGCAGAGGCT TCCAAGTCTA GTCTTGGCTC TATCACTTTA
30651 CGTGTTTATC CAGCTTGGTT GATCTTTCTG GACTCAGTTT CTATATCTGT
30701 AAAATAAGTG GTTTGGATCA GATGATCAAT AAAGTATCTT TTGATATTAA
30751 CATCGTAATA AATAGCTAAT ATTTCTTGAG TGCTTCCTAT GTNNNNNNNN
30801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNTGGAAGAT
30851 TATGTTCAGA AGACCATAAA AATTAAAATT TTTGTGGAGA ATAAAGTACT
30901 GATAATTCTA ATTGGCATGC ATAGTAATTT TATGGCCTCT GTGTATGTAA
30951 CCCACTGATC TCTTTATGTA AGAAGGACCC AGATTGACC ATAAATTTGT
31001 GTATTTTTTA TATTCTCACA ATAAAATAAT CTTGATATAT GGTTTTCTGT
31051 AATTTAAGAA AATATTATTC CTATGAGTTT CAATAATTAT TTCTAATGGA
31101 CATTAAATTT TAATGAAATT GACATCATTT ATAAGTCTGT TAATTAAGTT
31151 ATCGATTGAA AATTAGATTT GTGAACCTCC TGCCAAGTAG CTGTCTTTTG
31201 AAGATATTTT AGTATCTTTT AAACATTGTT TTTCAGATCA CAATTAATTT
31251 GAATGATGTA ACTTTTTAAA ATTCCAAACA AAAATAGCAC TTTTATTGTA
31301 AAAAATAACT CTTTACAGTT TATAACTAAA ATTTGAAAAT CTTAAATTTA
31351 TATGTAGTTC ATAAATGACC CTTTATTTAG GAGTCTCCTG CTTTCTACTT
31401 GCCTTTTAAC TAGATTGTTC TCGACTCCCA AAAAATTGAC TTAATTTTTT
31451 TACCATCTCC AACATGTTTT TATAGGGGCA CTGGCCGGAT TAATAGACAT
```

```
31501  TGCTGCCGAT  TGGATGACTG  ACCTAAAGGA  GGGCATTTGC  CTTAGTGCGT
31551  TGTGGTACAA  CCACGAACAG  TGCTGTTGGG  GATCTAATGA  AACAACATTT
31601  GAAGAGAGGG  ATAAATGTCC  ACAGTGGAAA  ACATGGGCAG  AATTAATCAT
31651  AGGTCAAGCA  GAGGTAAGTC  TTGCTTTGTC  TCAAGATGAA  TTAATAATTG
31701  ATATAGCAAA  ATGTTTCCAA  TTCATTTAAT  TATAGAACTA  ATCACATATT
31751  AGATGATTAC  ATACACATCA  AATGGATCCA  CCCTCAACAC  ATTGCAGCAA
31801  GAAAGAATTA  AGTGCAATAT  TGTTTCAAGT  AGCTTTTTTA  TTAGTTAACT
31851  GCATAGTCAT  ATAACAAATC  CTCTGGATTG  TGGTGCAAAT  ATATTTGAGC
31901  TGTAGTAGAA  AAGAAGTGAT  AGTTATTGCA  GTAAGATCTG  TGTAAAGTTA
31951  CTAAGAAGTC  AAGTTATTAA  AACTAATATA  TTACTAAGAT  TGGGAAGTTT
32001  GAATTATGAA  AGTATTATCA  AATAATTTAG  TAAAATCAAC  CTACGTAGAG
32051  ATACATTGAA  GATAATCAGA  CATTTTTATT  TGTGGCATTA  CAGCATTTAA
32101  ATGATTGATT  TACTATGATC  TACAAAGAAC  ATTTTAGAAC  TTAGGATGTT
32151  ACATGTATAT  TTTTTACATG  ATGACATGGA  TATATTTTTT  AAATTTTGTT
32201  TTAGCTGAAC  TTTAGAGCTA  AAAGGTATAC  ATTTGCGGTA  AGATGAGTAG
32251  TATGCTGTTT  CTCACCTGGC  TTAATTGAAT  TGAGTTTAAT  GATCTGGAAA
32301  GTTGCAGCAG  AATGAAATCT  GAGTGGTGAT  GCAATTGTT   TCCACTGTTT
32351  CCAAAAAGTG  GTTTGTAGGC  AGAGATTGAA  GTATAGCTGA  GATGTGTTGG
32401  TAACAAGACT  TTAGGGATTA  GGAAAAAGAT  TAAATGTGCT  CAGGGTTCCT
32451  TGGTATATGT  AGGCATTAAT  TTTTGGACTC  TACTTAAATA  TTTTGTTCAT
32501  ATAAAGTTTT  TATTATTGTG  GAAATAAACC  AGGAGACTTT  TACACATTTT
32551  ACTGAAGTTT  CTTTTCTTTC  TTTTTTTTTT  TTTTTTTTTT  TGGCCGGTGG
32601  GATGGAGTCT  CACTCTGTTG  CCCAGGCTGG  AGCGCAGTGG  CACGATCTCG
32651  GCTCCCTGCA  ACCTCCGCTT  CTGGGGTTTA  AGCGATTCTT  CTACCTCAGC
32701  CTCCCGAGTA  GCTGGTATTA  CAGGCGTGCG  CCACCATGCC  CAGCTAATTT
32751  TTGTATTTTT  AATAGCAACG  GGGTTTCACC  ACATTGGCCA  AGCTAGTCTC
32801  GAACTCCTGA  CCTCAGGTGA  TCCACCCGCC  TCAACCTCCC  CAGTGCTGGG
32851  ATTACAGGCG  TGAGCCACCA  TGCCTGGCCG  TTTACTGAAG  TTTCTTATGA
32901  CAAGCATTTG  CATTAGAGGT  GCAATGTAAA  TTAAATTCAT  ACTCTCGAAC
32951  TATTTTCTTT  TTAGGGTCCT  GGTTCTTATA  TCATGAACTA  CATAATGTAC
33001  ATCTTCTGGG  CCTTGAGTTT  TGCCTTTCTT  GCAGTTTCCC  TGGTAAAGGT
33051  ATTTGCTCCA  TATGCCTGTG  GCTCTGGAAT  TCCAGAGGTA  AGCCAAGTAA
33101  TATTTAGTGT  CATTAAACAT  TATTATGATG  CTTATCTTTT  TGACCTTAGT
33151  GATAATAAAA  GTTGGCTTTT  CTGGAGGGAG  GGGATAGTTT  GTTCATAATA
33201  TGAAAAAAAA  ATTTTTTTAA  GTATAAGCTG  ATGGTAGACA  TCATTGAAAA
33251  ATATTGTTCC  CCATAGTCAT  TTGGTCATTT  ACTGTGAAGG  CTGATTTTTT
33301  TTTTCTCTCA  CCACTAATTT  AACACATGAC  TAGGCAAATT  TTCAGACTAT
33351  TTAGTTAAAC  ATCAAGAGCC  TGGAAGAAGT  ATCTTGTGAC  CTAATGTTCT
33401  TTGACGGGTT  AGTTGTTACT  TTGCTGTTAT  GACCCTGAAT  TTTTTTTTTT
33451  TGAGACTGAG  TCTTGTGCTG  TCGCCCAGAC  TGGAGTGCAG  TGGCGCAATC
33501  TCAGCTCACT  GCAACCTCTG  CGTCCCAGGC  TCAAGCAATT  CTTGTGTCTC
33551  AGCCTCCTGA  GGAGTTGCGA  TTGCAGGCAC  CTGTCACCAT  GCCCTGCTAA
33601  TTTTTGCATT  TTTTTGTTTG  TTTTTTTTTT  TTAGTAGAGA  TGGGGTTTCA
33651  CCATGTTGGC  CAGGCTGGTC  TCAAACTCCT  AACCTCAAGT  GATCACCCGC
33701  CTCAGCCTCC  CAAAGTGCTG  GGATTACAGG  TGTGAGCCAC  CACACGTGGC
33751  TATGACCCTG  ATTTTGATTC  ATTCACTTTT  TATAATTACC  TTTTGATTAG
33801  ATAAGTTAAT  TATTCTTGAA  TTTGGCCATT  TTATGCTTTG  AGAAAGTAGT
33851  TAATCACAGT  GGGTCAACAG  TACAAACTTT  TGGGTTTTAT  TTTTCATCAC
33901  AATAAAGTAG  AGTTATACAT  AGGATTGATT  GAACTTGATT  TGAACTTATC
33951  TCTTCTCTTT  TATTTTTCTG  GAGTTAAATA  AGTTACCAAC  TTTTTCCTAA
34001  TACATTTCTT  TTTAAAATGG  AATTGTATTG  ATCCTTTAAG  TTTGTATTAA
34051  GAATATCTTT  CATAAAAAGC  AATATCATGC  AGTATATAAC  AGTTGTTACT
34101  CATTCTTGAT  ACATAAAAAA  CTATTGCACA  TAATTACAGG  ACCTCAGAGA
34151  AAACATAATA  TTCTTATTTC  TAACATAATG  GCCAAAATAT  ATTTAAAATA
34201  TTATGCTTAT  TTTTACAACA  GAAATATTCA  AATTTGCCCT  TTTTTTGGGT
34251  ATGTAATTAT  AATCCTTATA  ATTAAGGTCT  GTATTCATTT  TAACATGGCC
34301  TGATATTTTG  ATTTTGGCCT  GAGATAGTGT  TGCCCTCTCT  CCTTTCTTGG
34351  GTAGAGAATT  AGATTATAAT  ATCAATTTAT  TATATGTAGC  ATAATAGGCA
34401  AGTTTTCGAA  AAATTAACTG  TAAATTTTTC  TGTAGACTGC  TAAAATTTGC
34451  AAGGTTGTTT  TTGTGCATAA  AACAAGAAAA  TAACTTGGAT  TCGTTACATT
34501  CTCATGTTTC  TTAAAGGACA  TTAAGCTGCC  TTAATCTTTG  CCTTGTAGAT
34551  TAAAACTATT  TTAAGTGGAT  TCATCATCAG  AGGTTACTTG  GGAAAATGGA
34601  CTTTAATGAT  TAAAACCATC  ACATTAGTCC  TGGCTGTGGC  ATCAGGTTTG
```

```
34651 AGTTTAGGAA AAGAAGGTCC CCTGGTACAT GTTGCCTGTT GCTGCGGAAA
34701 TATCTTTTCC TACCTCTTTC CAAAGTATAG CACAAACGAA GCTAAAAAAA
34751 GGGAGGTAAG TGTCTTTTGT AGTTAATTTG ACTGAAAAAT ATATATTATA
34801 TAGTATTTAT TTAAGTAAAG AATTTCTTAG TGTAAAAATA ATAAATTCTG
34851 TATTCAGATA AAAAATTTTG AGATTTGTGC TTCTGTTTTT CCTGAATAAT
34901 CTATAACATC TTTCTAGAAT CCATTCCCAG TGCTGCTCAG TTCGTCTTAC
34951 ATTTTAGAGA AGCTTTAGAT AGACAGCTGG TGTCCATTGG GTTTCAGCTG
35001 CATTTCACGA AGATCTTCCT GTTATCACTT TACCTTACAT CTTTCCTCTT
35051 CTGAAGTGTT TTCTAAGCTT AGCTTTGTTT TTCACTCTTA CTTTCAACAT
35101 TAAGAGGTTG GGAAATCTTA ATAGCTATGT TTTCCTCCTG GAGGCAGTGT
35151 CTGGTGCCAG TGTAAGTGGT GTGTGATATG AAAAATGCTA TCCAGTGCTA
35201 TGGGGAAGTT CTGAGGGCCT TTAGAAGCTC TTGAAGTTTA AATCAGAAAT
35251 TCACATTAAA GAGATTACAG GAAATCCTTT TCATTTGATT GTTTAAGGCA
35301 ATTTCCTTTA CCATTTCTTT AGGCCAGCCT GAGATCTTCT ACAAGACCTT
35351 GAAACCTTAT ATATATTATG GATTTCCTCT GATGTTTCCA TATTGCTCTG
35401 GGCATTTTCC TGAATCCTTT ATATTAGCTC TAGACTTTGG GAGCCCAGTC
35451 CCTTCCTATT TTCCAAATCT AAATCTACAG CCCTAGATGG TACAGAGATC
35501 TTTGAGTTTT TAAGATATGA TTTTTTGAAA AACATCTCAT TAAATACTGG
35551 CAGAACCTTT TCATCTTGTT GAGTTTTTTA ATGTACTGTA ACCAAAAAAG
35601 TAGAATATTT TATCAAACTG TTTAATCTTC AATTGAAATA ATTCTAGTAC
35651 ATTTTAATGT TCGCATTAAA ATATTGTCCT TGCATTGGAC GTAGATATCC
35701 CAAAAGTGGA ATACTTCAGA TTGTCGTAGT TTCATCTCTG AATAATTGTG
35751 ATTCCAGTAC TTTATAACAA AAATAGCTAG CATTATTGAT TACTTTCTGT
35801 GTATCTGGTA CTGTGGCAGA TACTTTACTT GGATTTAAT ACTTAATTTC
35851 ACAGTAATTT AGTAATATGG CCCTGTTATC CTCATTTAGT GATTAGTAAA
35901 CTAGGGCTGA AAACAGCTAA CTAACTTGCC CGAGACTACA TACCTAGTAA
35951 GTGGTGGAAC GTAGGTTAAA ATTCATTTTT CTTTGACTTC AAAGTCTGTG
36001 GTCTTACCTA CTTACATTAC TGCCCTTACG ACTATGTGGG TATATATTTG
36051 TGTGTGTTCA AAACAAACTC AAAACCATCC TGTAGCGTAG CAAGTTAGTG
36101 GCTAAGATGA AGCTAGAGCA TTTGCCTCCT CAATTCAATT CCATTACTTT
36151 CTGTTGTACC TTTATATTTT TTGGTAAGAC TTTTACTTAT TCTAAGTTCA
36201 AAAAATGTAA TTTATTAGAT GTTTGAGAAA TTAAGTTTAC CTAAATTTTA
36251 ATGTTCATAC TGTAGTGATT AGTTAATGTT TAATACGTTG TTATTCTGTC
36301 ACCTTAGTGT ATATATAAAT GGCAAGAATT CACGGTTAGT TGAAAGCATT
36351 AAGGTCCCAT AGTTTTGTGT AGACAAGAGG GGAGAGCGTT GATATTTTTA
36401 AATTAAATGC TTCTTAGATA CGTATGAAAT GGATTAAAAC ATGTATATGA
36451 GTTATAGATA CCTAGGTGTT AGTTGGTTG TAAATTCAGG ATCAGGACAT
36501 TCAAATAAAT ATGTTTGCTT TCCTCTTAGT GGAGGAAAAA AAAAAGAAGC
36551 TAAATTTGCT CCCTTTCCTC CCCAAATAAG CAGAGTCTAC ATTTTAATGC
36601 CAACAATTTG ATTAAAACAA ATATTTATTT ATTTTTAATT CACCAAACTT
36651 TTATAAAGTA TTTACTGGTG CCAGGCACTG TTCTAAAGCA CTCTGTATAT
36701 ATTTACTCAG TCCTTAAGAG CTAAGTAATA TTATCACGTT TCCATTTTAG
36751 AGAAACTGA GGCACATATA GGTTAGGTTA TCTACCCATA GCCATACAGC
36801 TAGTAAGTAG CAGAGCCATG ATTTCAACAC AGCAGCCTGA CTATGGAGTT
36851 CATGATCTTA ACCATTTACA GCTTAATTTT TATTATTTAT AATTTCTCTT
36901 CTGGAAATGT AACAATTGAC CATTTGAAGA AATACTTTAG GTAGCTTTGG
36951 ATATTTGCTG TATTAAAGTA GTGAAAGTAA TACAGACACT TGGCTGGGCG
37001 CGGTGGCTCA CGCCTATAAT CCCAGCATTT TGGTAGGTTG AGGCAGGCAG
37051 ATCACCTAAG GTCAGGAATT CGAGACCAGT GTTGCCAACA TGGTGAAACC
37101 CCGTCTCTAC TAAAAATACA AAAATTAGCC GGGCGTGGTG GCAGGCGCCT
37151 GTAATCCCCA GCTACTCGGG AGGCTGAGGC AGGAGAATCA CTTGAACCCA
37201 GGAGGTGGAG GTTGCAGTGA GCTGAGACGA CGCCATTGCA CTCCAGCCTG
37251 AGAAACAAGA GAGAAACTCT GTCTCAAAAA AAATAAAGGA ATACAGACTC
37301 TTAGAAAAAT AATTACAAAT AAAACCCTAG TGAAATTATA GGTATAGTTA
37351 GGTATAGTTG GCTTACAGGT GGGAAGTAGA CCATTACCAA CTGATAGACT
37401 GGGGAGCTGG AGAGAGGACA CGGAAGAGTG TCCTTGGATT TTTCNNNNNN
37451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
37651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NAAAATTGTC TATATTCATT
37701 GCCTCCTCCT CTTTACACCC TATTCACATT AGTATATCTG GCAAAAATTT
37751 TTTTTAACTG AATGGTAAAT GCATGACTGA CCTTTCAATT AAAGCCAGGA
```

```
37801 GAAAGAAACA AATCTTAATA GAAGAAATGA ATAGTTACCC TTTGCTTAGG
37851 GAGCAAGGAA ACATGCAAGT TAAATTCAGA AAATCCATTT GGAAAATTCA
37901 AGTAACATGA AGAATTTTTA TTTGGTATGT TTGAATTTCT ATGAAATTAT
37951 GAAATAAGCC ATATCCTCTT TCTAGGTGCT ATCAGCTGCC TCAGCTGCAG
38001 GGGTTTCTGT AGCTTTTGGT GCACCAATTG GAGGAGTTCT TTTTAGCCTG
38051 GAAGAGGTAG GTGAAAAGAA TACAACAATT AAAATTATAT ATAATTACCA
38101 TTACAAATAT ATTTCACACA TTTCAGTTTT GTAGGTGATG TAATAGGTAG
38151 AGACTTTGTT TTCAAATTTA TTTTTCTAAA GTTGTTTTCC ACTCATTCTT
38201 AATAAAAAGT AAATGTTATT CATGCTCCAT ACCTGGAGGA AACTTTTTAA
38251 AAATTTATTA ATGTATGAAT GTTAGTAATT ATTTAAAATC TAACTTTGTT
38301 GACATATTTA AAAGTAAGAA GATGTGAATT TGACTTAATA GAGGACATGT
38351 GAAACAATCT ATTTCCATTG GCTAAATTCT GTATTTTTAG TAGAGATGGA
38401 ATTTCACCAT GTTGGCCAGG CTGTTTTTTT GTGGGGTTTT TTTGTTTTGT
38451 TTTGTTTTGT TTTTGTTTTT GAGACGGAGT TTCACTCTTG TTGCCCAGGC
38501 TGGAGTGCAA TGGCGCGATC TTGGCTCACT GCAACCTCCG CCTCCAGGGT
38551 TCAAGTGATT CTCCTGCCTC AGCCTCCCAA GTAGTTTTTG TTTAAAAAAT
38601 TTTAATCAAT TCCTATGTTG AGTTTTAAAG TTTTTCCCAT GTGATTATTT
38651 CTGATACAGT TAGTGATGTT AAAGAAAATA ATTTTAGTGA CTTCAGTGGA
38701 TTATTTTGTT TTTGTTTTCT TAATAGGTGT TTAAGACTTT TCTTTTTACA
38751 TAAAAATGTA ACCAGGAATT TTTTTTTAAT TTTTTTTGAC AAATAATAAT
38801 TGTTTTTGTT TATGGGGTAT AATGTGATGT GTCTATACAT GTATACATTG
38851 CGGAATAATC AAATCAGAGT GATTAGCAAA TCCCTCAAAT ATTTATTATG
38901 TCCTTGTGGT GGTGAGAACA TTTAAAATCC TCTTTTAGCT ATTTTGAAAT
38951 ATATAATACA TATTATTAAC TGTGGTCATC TTACTGTGCA ATAGAACACC
39001 AGAACTTATT CCTCCTCTGT AAGTTCATAC CCGTTGACTA ATGTCTCCCC
39051 TTTCCCTGTT CACCTCCCCA ACCCCTAGCC TCTGGTAACC CCTATTCTAC
39101 TCTCTACTTC TATGAATTTA ACTCTTTTAG TTCAAGATGT TTTTAAATGT
39151 ACTTTTTTCT TTTAGTTGTT TGTATTCTTT TTTTTTTTTT AATGTAGAAG
39201 AGGCAAATTA AATGCATTAT AAGTTAACAG GAGTTGGTGA TGGTACATTT
39251 ATTTTTAACT ACCATGATTG AATTGAATGT GAAACTCATT TTGAATATAA
39301 AACAGCACTA GGTATTCTAT TAGTATTTAT TAGACATTTA TGATCAATTG
39351 ATACTGTCAA TTTGTAATGA TGATCACCAT CTCCAAAAAT AATAATAACA
39401 TCAATTTTTC TTATTACAGT AAAATCCATT ACATGTAAAT TCTAACTACA
39451 GCAAAATTTA GAGCTAGGAT ATTTACCATT CAAGTTATAA TATATCAGAA
39501 ACATCTTATA AAATTATAGC ATTAATTTTT CTTTTCCTTT TCTTTTTTTT
39551 AGGTTAGCTA TTATTTTCCT CTCAAAACTT TATGGAGATC ATTTTTTGCT
39601 GCTTTAGTGG CTGCATTTGT TTTGAGGTCC ATCAATCCAT TTGGTAACAG
39651 CCGTCTGGTC CTTTTTTATG TGGAGTATCA TACACCATGG TACCTTTTTG
39701 AACTGTTTCC TTTTATTCTT CTAGGGGTAT TTGGAGGGCT TTGGGGAGCC
39751 TTTTTCATTA GGGCAAATAT TGCCTGGTGT CGTCGACGCA AGTCCACGAA
39801 ATTTGGAAAG TATCCCGTTC TGGAAGTCAT TATTGTTGCA GCCATTACTG
39851 CTGTGATAGC CTTCCCTAAT CCATACACTA GGCTAAACAC CAGTGAACTG
39901 ATCAAAGAGC TTTTTACAGA CTGTGGTCCC CTGGAATCCT CTTCTCTTTG
39951 TGACTACAGA AATGACATGA ATGCCAGTAA AATTGTCGAT GACATTCCTG
40001 ATCGTCCAGC AGGCATTGGA GTATATTCAG CTATATGGCA GTTATGCCTG
40051 GCACTCATAT TTAAAATCAT AATGACAGTA TTCACTTTTG GCATCAAGGT
40101 AAGTGCTAAT GTGAGGTGAT ATTTGGGTAA TTTTGGCATG TTCAAAACTT
40151 ATATGTGGAA TGAGAGAGGT TGTTGTTTCA TAAATGACTG AAAAAAGTAC
40201 TTATCTTTTG AGTTTAATTT TAAGTAATGA AAAAGATAAT TCCTTAGCAT
40251 ATATTGTTGA CCATGTTATC TGTTGCTATT TAACAAATTA CCCCCCAAAA
40301 CTTAGCAGCT TAAGGTAACT ACTTATTTTG TTCTTGATAT TGAGTCAACG
40351 ACTTGGGAAG GGCTCAACTG GGCAATTTTT GCTTGTGGTC TTTCATATAG
40401 TTGTTATTAG ACATGGCGAG GGCTAATCAT CTCAAAGCTT CTTTTTTTCG
40451 TTTCCTTTTT AAAAAACTGT TTTTGTGGAT ACACAGTAGC TATATATAGT
40501 TTTGGGGTAT ATGAAGTATT TTGATAGAGG CATGGAGTGC ATAATAATCT
40551 CAGGGTAAAT GGAGTATCCA TCACCTCAAG CATTTATCCC TTGTGTTACA
40601 AACAATCCAA TTACACTCTT AATTATTTTT AAGTGTACAA TTAAATTATT
40651 GAATATAGTT CAAAGACTTC TTCATTCATG ACTAGCACCT AGGCTAAAAA
40701 AATTCAGACA CCTGGGCTCC TGGGATCAAT CACGCATACT GTGTCTCTTG
40751 TGCTCACTCC CGCTGTCTCT CTCTCTTTCT CTCGCTTCCT TTTTCCTCTC
40801 TCTCTGTGGT TTTCTAGGGT GGTGGCCTCA GGGAATTGGA TTTCTTATAT
40851 TATAGCTCAG GATTCCCAAG AGGGCTGTTT TTAATGTAGC CAAAGAAGTC
40901 TTGCAGCGTG ACTTGTTTTA TTCTATTCAT TGAGGTAGTC ACAGAGGCCC
```

```
40951 GACCACATTC AGAGGAGGGA CATACACTTG CTGGGACAAG TGTAAGAGAA
41001 TTCATGATCA TGTTTTAAAA CCACTTTTAT TAGTTTCCTA TTGCTGCTGT
41051 AATAAATTAC CACAACTTAA TGGCTTAAAA GCCACACAAA TTTAATATCT
41101 TACAGTTCTG CAAATCAAAA GTCTGAAACG GATCTCACTG TGCTAAAATT
41151 AAGGTGTTCG TAGGGCATTC TGGAGGCTGT AGGAGAGAGT CTTGTTTTTT
41201 GCCTTTTCTG GCTATTAAAA GCTGCCAGCA TTCCTTGGCT CCTGGCTGTC
41251 TATTTGCATC TTCAAAGCCA GCAGTAGCTG GTCAAGTCTT TCTCTTGTCT
41301 CATCACCCTG ACCCAAACTC TGCTAAATCT CCCTTCCACA TTTGAAAAAC
41351 CTTTGTGATT ACTTTAGGCC CACGCAGATA AATCAGAAAA TAATCTCCTT
41401 TTTCAAGGTC AGTTGCTTCG AAACTTTCTT TCTGCCACCT TGATTCCTCC
41451 TTGCCATGCA ACGTAATGTA ATCACAGGTT CTGGGAATTA AGTTATGGAC
41501 ATCTTTGATG AGCCATTATT CTGCCTCATA CCAGTATAGG GTATTAGCTT
41551 GAAAGGACAC TGCAGACTCA GTTAAATTAC TAGATCTATA AATACATGCC
41601 TTTTTCCATC AAGAAATTAA GGCAGCTGGG TCTTATGCCC TGGGACATTG
41651 CTTCTTTTGG ATTTATAAAA TAACAAAATT TGTTGATTAA TGGTCTATCA
41701 GTAAATATAA TTTCTTATGT GACTATCAGT GATATATATG GGGAAGCACA
41751 TATCAGCTTA TTCTTGTTCT TTAAATTACT ACCCCTGTAC TTCATGTAAT
41801 AGTATTTGCT AGTGATGATG TGCTTTTACA GATGTAAATT AATGTGGAAT
41851 AACAGCTTTG TTTCTACAAA ATTAGAGTGG TTTTAGTTTT TGAAATAAGG
41901 TCTCTTTTCT CTTGTCCTAA GTCTGTAGTC CACTGAGTAT CTAGAGTTAA
41951 ATAATAGAAA AGCCTGGCCA GGCGCAGTGG CTCACACCTG TAATCCCAGC
42001 TCTTTGGGAG GCCGAGGCGG GCAGATCACA ATGTCAGGAG ATCGAGACCA
42051 TCCTGGCTAA TGCGGTGAAA CCCCGTCTTT ACTAAAAATA CAAAAATTAG
42101 CCAGGCGTGG TGGCAGGTGT CTGTAATCCC GGCTACTCGA GAGGCTGAGG
42151 CAAGAGAATC ACTTAAACCC AGGAGGTGGA GGTTGCAATG AGCCAAGATC
42201 ACACCCACTG CACTCCAGCC CAGGCAACAG GGCAAGACAC TGTCTCAAAA
42251 AATAATAATA AGAAGAAAAT AATAATAGTA ATAGAAAAGC CTAAACATTT
42301 TACCTTTTTT TCTTAGGGAA TCAAGTTAAA AGAGCTGTTA AAGCTCTTTT
42351 TCCTACAATA AGTAAGTGTT GGGTAAATCC CAACTTTCTC ACAGTCAGTT
42401 GAACTACAAG AAGCTGGAGG CAATGGCAG GCCTTTGTTA AGTCCCACCT
42451 TTGACTCAGC TCTGGCTGAA GGATCATACC TGGCAAGAGA GTGTAAAACA
42501 CACTTTGATT TTTTCTATTG TTTATCCTTT TAATGATCCT AAGAGACTCA
42551 AGAGTACATG CCATCATTTT GTGTTTGGCT CATTTCATAT TCAGAGGAGT
42601 TTATTACTCT TTCAGTAGTT TGTTTGTTCG TTTGTTTGTT TTTTGAGACA
42651 GGATCTCGCC TTTTTGCCCA GACTAGAGGG CAGTGTTGCA GTCTTGGCTC
42701 ACTGTAACCT CCACCTCCCA GGTTCAAGCG ATTCTCCTGC CTCAGCCTCC
42751 CAAGTAGCTG GGATTACAGG TGTGGGCCAT CACACCCGGC TAATTTTTGT
42801 GTTTTTAGTA GAGATGTGAT TTTGCCATGT TGGCCAGGCT GGTCTGGAAC
42851 TCCTGACCTC AGGTGATCCT TTGGGAGGCC TTGGCCTCCC AGAGTGCTAG
42901 GATTATAGGT GTGAGCCACT GAACCTGGCC TCTTTCAGTA GTCTTTAAAT
42951 GATCTTGCTT ATGGTGCTTC TTATCCCTGT TTATTATCCT TATTAAATTT
43001 AATCAATAAA TATTTTTCTC TTTTTAATTG ATTCATATAA ATAGACTTAC
43051 CTGAGAGATA TAGGTTCAGT TCAGAGCACC ACAATAAAGT GAATATCATA
43101 ATAAAGCAAG TCACATAAAA GTCTTAGTTT CTTAGTGCAT ATAAAAGTTC
43151 TGTTTACACT ATGCTGTAGT CTTATGTGTA CAATAGCATT ATGTCTTTTA
43201 AAAAAGTAAT ACTTTAATTT AAAAATACTT GATTGCTAAA AAATGCTAAT
43251 AGTAATCTGA GTCTTCAGTG AATTGTAATC TGTTTTGCTT CTGTAGGGTC
43301 TTGCCTTGAT ATTGGTGGTT GCTAGAGGTA GGACTGGCTG TAGCAATTCT
43351 TAAAATAAGA TAACAGTGAA ATTTGCCGCA TTGATTGACA CTGCCTTTCA
43401 TGAAAGATTT CTCTGTAGCA TGTGATGCTG TTTGATACCA TTTTACCTAC
43451 AGTAGACCTT CTTTTCAAAA TTAGAGTCAT CCTCTCAAAC CCTGCTACTG
43501 CTTTATCAAC TAAGTTTAAG GAAAATTCAA AATCTTTTGT CCTTTTAACA
43551 ATGTTCACAA CATCTTTACC AGGACTGGAT TCTACCTCAA GAAACCACTT
43601 TCTTTGCTCA TCCATAAGAA GTAACTCCTT ATACATTCAA GTTTTTTAAA
43651 TGAGATTCTA GCAATTCAGT CACATCTTTA GGCTACGCTT ATCATTCTAG
43701 TTCTCTTGCT ATTTCCACCA CTCTGTAGTT ACTTCTTCAA CTGAAGTCTT
43751 GAACCCCTCA GAGTCATTCA TGAGAGTTGG AATCAACTTC TTCCAAACTC
43801 CTGTTAATAT TGATATTTTG ACCTCCTCCC ATGAAACGTG AATGTTCTGG
43851 ATGGCATCTA GAATGGTGAC TACTTTTTGA ACATTTTCAA TTTATTTTGC
43901 CCGGATCAAT CAGAGAAGTT GTTATCAGTG GTGGGTTTCC AAGTTGTCAG
43951 GGGCGAACCA TACAGATCTT CAGCAACCTC AACTCTTGCC TTCTCAGAGG
44001 AAAGAATTCT ACGGAGGGAC ATAAGGCAGA AAAAGAGACT GAGGCAAGTT
44051 TTAGAGCAGG AGTGAAAGTT TATTATTAAA AAGCTTTAGA GTGGGAATGA
```

FIGURE 3, page 14 of 27

```
44101 AAAGAAATTA AAATACACTT GAAAGAGGGC CAAGTGGGCA TCTTGGAAGA
44151 CAAGTGCCCC ATTTGACCTT GGACTTAGGG TTTTATATGT TGGCATACTT
44201 CTGGCATCTT GCATCCCTAT TCCATTGATT CTTCTTTTGG GGTGAGTTGC
44251 CCACATGCTC AGTGGCCTGC TAGCACTTGG GAGGGGAGTG TGCACAGTGT
44301 ATTTACTGGA GTTGTATGCA TGCTTACCTG AGGTGTTTGT TGCTTACCAG
44351 CCAAATGTCC CTAGGAGGTC ATATTCATAA ACTCCATGAT TTTGCCTCTA
44401 AATGTGCATG CTTGAGCCCA CTCACCCAAC TCCTGGGATC TTATCGGAAA
44451 GCTGCCGATC GCTAGTTTCA GGTGTTTCTA TCTATTGGAA GATGGCCTTT
44501 CCCTGATGCT GGCTGCAACC AATTATTACT TTAGAGAGAG AGCATGAGAG
44551 CTGTCTCACC ATCATCACCT GATGGTTGCC TGACATTCCT GGTGGGGTTG
44601 GGAGGATGCC TGTCCTGCCC TGCTCATGCC TGACTAGCTA CCTGCTGTAA
44651 CAAAAGTACT ATCTATGGTA GCTGTAGCCA TAGGAAATGC ATTTCTTCAG
44701 TAAAACTTAA AAGTCAAAAT TAGTCTTTAA AACAACATGA ATCTCCTTGT
44751 ACATCTCCAT CAGAGCTCTT GGAAGACCAG GTGCATTATT AGTGATGAGT
44801 AATGTTTTAA AAGGAATCTT TTTGTCTGAG CAGTAGGTCT CAACAGTGGG
44851 CTTAAAATAG TTAGTAAACC ATGCTGTAAA CAGATATGCT GTTATCCAGG
44901 CTTTGTTATT CCATTTATAG AGCACAGAGA GAGTAGATTG GCATAATTTA
44951 AGGATTACTT AAAAAAAAAG TCTTTGATTA CTCTCAAAAA AAAGTCACGT
45001 CTCTCACTTT ATATCAACAG CTAAAAATGG CCAGGTATTG TGGCTCACGC
45051 CTGTAATCTC CATGCTTTGG GAGGCCAAGG CAGAAGGATC ACTTGAGGTC
45101 AGGAGTTAGA GACTAACCTG GGCAACATAG TAAGACCCAT CTCTACAAAA
45151 AAAAAAAAAA AAAAAGAAA GCCAGGTGTG GTGGTGCACG CCTGTAGTCC
45201 CAGCTACTCA CGAGGCTGAG TCGGCAGGAT CACGCCCAGC CAAGAGACGT
45251 GACTTCTGCT TTCAGTTGTA CACTTAGAGA CCATTGTAGG GTTCTTAGTT
45301 GGACTAATTT CAATATCATT GGGTCTCAGG GAATAGGGAA GCCTGAGAAG
45351 AGGGAGAGAC AGGGGAACAG CCAGTTAGTG GAGCAGTCAG ACCACATACA
45401 ACACTTATTA AGTTCACTTT CTTCTATGGG CATGGTTCAT GGTGCAGTAA
45451 AACAACTGTA ACAGGAACAT CAAAGATCAT TAATCACAGA GCACTGTAAC
45501 ATATAATAAT AGTGAAAATT TTCAAAGTAT TGAGAGAATT AGCAAAATAT
45551 GATACAGAGA CACAAAGTGA CCACATGCTG TTGGAAAAGT AGTGCTGATG
45601 GACTAGCTTG ATGCAAGGAT GTCATAAACC TCAATTTGTG AAAACTGCAA
45651 CATGTGTGAA GCACAGTAAC ACAAAGCATA GTAAAACAAG ATATGTCTGT
45701 ATATCAGTCA AAATATTGGG CAACTCTGAT AAGTTTGTCC ACTTAACATT
45751 GTACCACTTA AGATGAATAG CATCTACCAT TTCCGTCATT TGTAAATATA
45801 TAGGAGGACA TAATCACATA ATCTTGAAGT AAAAGACAGT GCTTAAAACT
45851 GAATCAGTTA AGTTTTATGA AAAATACTTC ATATTGTACT TTTAAAAATA
45901 TATATTTTTT AATTTCAATA GCTTTTGGGT TACAAGTGGT TTTGGTTACG
45951 TGGATGAATT CTATAATGGT GAAGTCTAAG ATTTTACTGC AACTGTCACC
46001 CAAGTAGTAT ATATTGTATC CAGCATATTG TCCTTTTTTT TTTCTTTTTT
46051 TTTTTTCATT TCACCATGGA CTAATGAAAA TTTTGTTAGG GACTGACATT
46101 AGGGCACCCT TGAGCTACCT TGAGCTAAAG GAAATAACCC TTGAATTTTT
46151 TTCTGTTTGG CCTAGAGAAT GTGGTTTGTT TTGTAACTGA ATTCATGGGA
46201 TTGTTAAGGT ACAAGATTTT GCTTTAGTTT TATTTGTACT AGGATTTTGC
46251 TATATTAATA CAATGTGAAA AGAATCAAAA GTGTTAGAAA TAAATGCATA
46301 GAATGTAAGT TTCAGGCATG TGAGTAGAGG ATCTCTGCTC CATAAAGAGT
46351 TCTGTTGTTG TTATAGGTTC CATCAGGCTT GTTCATCCCC AGCATGGCCA
46401 TTGGAGCGAT CGCAGGAAGG ATTGTGGGGA TTGCGGTGGA GCAGCTTGCC
46451 TACTATCACC ACGACTGGTT TATCTTTAAG GAGTGGTGTG AGGTCGGGGC
46501 TGATTGCATT ACACCTGGCC TTTATGCCAT GGTTGGTGCT GCTGCATGCT
46551 TAGGTAATAT GGCTGTGTCT GCCTGTGTGT GGATGTTGC AAGTCTGAGA
46601 GAGCCAAGAG AAAGTGGGAC ACATTCTTGC TTAATTGGTG GGCGGATTGG
46651 TTGAGTAAAG GAGGGTGCCA GGAGGAGATG TTTTAACAGA TAAGAAACAG
46701 TAGTACTATT AGGGTATTAT ACAGTACCGG TTTTCTGTCT TACAACATTT
46751 GTTAATACAA GAATTTAATG GCATTAGCAT ATTGTAATAT AACTTAATAC
46801 ACTATGGCAG AAGCCATCTA AGTACAACAT AAGCTTAATT TGAATCCTGA
46851 CCAAAGATGT CTTTGATTCT TTCATCGTTA AGGATCTTGG CTTACCTATA
46901 ACAACTATAG CATAATACCT AAGATTAGCA TTGCAACAGA GTTTCAGAGT
46951 AGGTTTACTT TGGTTCTGAA ATGATTTATT GTTAGCCTTA GTAAAAGATG
47001 TATTTACCCA TGCTCCATCA TCTAAGGTAT ATTTGTAACA AAATGAGAAA
47051 AGGTAACTTC ATTTTAATGA GAAGAAAAGC AAAATACCTA CATTAAGTAC
47101 TTGAGTCTAT TTAATGTCTG TTAGGGCAGG AAAAAATGGT TATTGCTTTT
47151 CATATTTAAA ATATCAGCTA CACTCTGGTG ATAATATTAA TGGTTGCCAT
47201 TTTGACCAGT TTTGTTTAGT GAATAAAAAT TATGTGATTA TTGATCTTTA
```

FIGURE 3, page 15 of 27

```
47251 AAAATGTAAT ATCAATTAAA AGGAAAGGAC AGACTCATTT TCACCAAAGT
47301 AGCAAGTATT TATTAAATGT CCACTTTCTT TTTAGCATTG TGCTAGATAC
47351 AGTGCATAAT ACAAAAAGAA CATGGACCCA ATCTCGACTC TAATCAAGTT
47401 GAGGAGACAA GATGAACACT GAGAATACAA TAGTGAGGAA TACTAACAAA
47451 TATATACAAG GTTAAAAGAG TCTAAGTATG GTAGGAATAT AGGGGAAGAA
47501 AGAGCTGAAG TACTTCAGGA AGAGTAGAAC ATGAGGCTTT ATTTAAAAGA
47551 TTAGCAGAAT TTAAGGAAAA GGTGACTTTG TTGAAGATTA TAATGTGAAG
47601 ACAAAGGAAC GAGGATGGGA ATAAATTTTG TATTCATGAG GCTTTGAAGA
47651 AATTGACTCT AGAGAGTATA TTTTGGGTAC TTTTGGGAAA TGAAGTTGGA
47701 TTAGTGAGAA GGAACAGATT ATGAAAAGAC AAGAAACCTG ATTAATGTCA
47751 GGATGATTTT ATATTTGAAG TTGGTCAGAT TTATGGCAGT CCTGGCTTTG
47801 CCATTTTTAG TTTGATGACT TTGAGAAAGT TCCTTCTTGA AGTTTTAATT
47851 TTCTGTATAT AAAAAGTAAT AACACCTGGT GATCTGCTAG GTTTGTTTTG
47901 AGGATTATAT GAGATAAAAT GCATGCAAAA CTGTTATAAT AGTGCCTGGT
47951 AAAATAAGTG CCTAGTTTTA AAAACAAGTC TTTGTAAACT GCTTAGGACA
48001 TGCCTGGTAT AGGGTAGGTA TGTAATACAT AGTAGGTAGG ATCTGTCTCC
48051 TTGCTATTTT TAGGTAAAAA AACAAAAGGA AGAGCTTCAG CTTAATACAG
48101 TATGAACTGA CGAGCCCTGG TAGGTTTTTG AGCAAAAGAG CAACACAGTA
48151 AAAGTAGTAC TTAGGAAAGA TTAACAAGGG AACATGGCTT ATACAGTGGT
48201 AATGGGGCCT GGAGTCAAGG AGGTAAGATA AAATGGTATT ATAATTAAGG
48251 AATAGCCAGG CACGATGGCA CATGCATGTA ATGCCAGCTA CTGGAGAGGC
48301 TGAGGTGGGA GGATCATGGG AGTCCAGGAG TTTGAGACCA GCCTGGGCAA
48351 CTGAGTGAGA CCCCAAATCC TAAAAAATAC AAAGTAAAAA AGGAATAAAG
48401 TCATGAGGGC TTGGACTGGA TTGATAACAG TGAGAATACC GAGAAAGGGA
48451 CCATAGGCAG TGTGAACGCA GCTCACTGCA GCCTCAAACC CCAGCCCAAA
48501 CGAGCCTCCC ACCTCAGCCT CCCAAGTAGC TGGGACCACA GACATACACC
48551 ACCATGCATG ACTACTTTTT TTAGTTTTTA CTTTTGTAGA GACAGGGTCT
48601 CACTGTATTG CCCAGGCTGG TCTCAAACTC CTTGACTTAA GTGATCTTCC
48651 TGCCTTGGCC TCCCAAAGTG ATTACAGGCA TGAGCCACAG TGCCTGGCCC
48701 AAATAGTTTT CTGTGAGTGA ATATTACTTG CATCGTTAAT GTAAATCAAA
48751 GGCATCAAAG TATTTTACTC TTTTTGAAAA AAATTTAGAG GAGAAATTTA
48801 TTATATTAAT ATTCTACCCA TATATGAGTT TAATTTGTAA ATTGTAGCAA
48851 AGCATGATGT GCTTTACTAA ATTCCTTTAT AATTAGAATA AGCTTTTATA
48901 AGGGTGAAAT TATGTCTTTG CTACAGCACT AAACCAAAAT GGCAAAATTG
48951 TTTTAGTCGG TAAGCTTTGC TTTTTTAAAA TATGAAATAA ACAGGTTTTT
49001 AAAATGTTAT TTTAATAGTC TTCTCTGTTA TAAACAAAGA AAATTGGTGT
49051 TTCTCTAGAG CTTATTAAAA GTAGTGATTA TTGTCCTAAA AGAGGAGTAG
49101 CAGTTTTAGA TGCTAATGCT TTTCCCTGAC TGAGTTCTAT TTGCCATTTA
49151 GTTTTAACTG CCTAGTGCAA AAATTCTAAT AAAATGTAAT GATGAGGATC
49201 CTGTCCTTCC TGACCAGTGG GTGCTTACTT TTTTCAGGTG GTGTGACAAG
49251 AATGACTGTC TCCCTGGTGG TTATTGTTTT TGAGCTTACT GGAGGCTTGG
49301 AATATATTGT TCCCCTTATG GCTGCAGTCA TGACCAGTAA ATGGGTTGGA
49351 GATGCCTTTG GCAGGGAAGG CATTTATGAA GCACACATCC GATTAAATGG
49401 ATACCCTTTC TTGGATGCAA AAGAAGAATT CACTCATACC ACCCTGGCTG
49451 CTGACGTTAT GAGACCTCTA AGGAATGATC CTCCCTTAGC TGTCCTGACA
49501 CAGGACAATA TGACAGTGGA TGATATAGAA AACATGATTA ATGAAACCAG
49551 CTACAATGGA TTTCCTGTCA TAATGTCAAA AGAATCTCAG AGATTAGTGG
49601 GATTTGCCCT CAGAAGAGAC CTGACAATTG CAATAGGTAC CCTTTCAAAA
49651 ATATATATAT GTATATATGA GATGGATTTC TGGAAGAAAG GAAAGCAATA
49701 AGCAGTAACA TTTAATGGGT CGGATTTGTG GGGGCAAGGG ACATTATTTC
49751 ATGTCCCTTA ACATCTTCTG TTCTTTAAGA AAGGAAGGTA TGCTTCAGTG
49801 GATGATTTTC TGCTATATAT CACAAAATCT GTATTTCAGG TTTGTCTTTT
49851 GATCCGGCAT GTACCAGAAA TTGGAGTCAG ATTATTTTCC CACTCAGATA
49901 AGCCTAGATA AGTTGATCTT GGTTATTCAA AACAGCATGT AATATAAGAC
49951 CTTAGCTAAA TGCATTCAGT CAAATACATT CTTGTATTTA ATAAAGTTGG
50001 CTTATTGGAA TACAAGTTAT TGAAAATCTC ATCTTCATCA GTCTCTTTCA
50051 TATTAGAATA ACACTGTTTT GCTTTATCAG TCTTTGGGGT TAGAATTATA
50101 ATATTAATTT ATAATATCTG ATTTAAAGTG ACAATCACTG AGATTTTTAT
50151 TTCTGATCAA ATGCCAGGTT GAAAAAGTAT AACGTATCAG TCCTGTTGTG
50201 TTTTATGCAG ACTTTCCTGA AAATACTGTT TAAAGGTATT AGCCATAGTG
50251 TATTTCTTGG AGATAAATTA AACTTTCTAT AGTTCTGTTT CTCTAAAATT
50301 TGTTTTTCTC TTTACCTTAT AGTCCCGCAG TATTGATGAG GAGACCATTA
50351 AGACTTAATA TTTTTTTTGAC ACAATCTTAT ATCTCTTCCT CCAACCCCTA
```

FIGURE 3, page 16 of 27

```
50401 AAAAGTGACT GAGGATAGGT ACATCAAGCC ATTGCTTTGT TACTCCCCAG
50451 GTTTTAGTGC CAGACCCTGA ATGGAAGTGT CAAGCCTTTG GCCTGTCTGA
50501 AAGGTCATTC CTGTGAGCAT ATCATCTCCC TTCCAGCTTA CCTCTGTGGC
50551 CATTGCAAAA GGATTTAAAA ATAATTTTTG TGCCATTTGA ATGGCACAAG
50601 ACCAGACAGT GTATGTGGGG GAGTGTTTCT CAAATCAAAC TGGAAACTCT
50651 TTAATTTGTA AGAACCATTA AGCAGAGAGA GAAAAAAGAA AGGAAAAGAA
50701 AAAAGATCCT ACAGAGAACA CCCTGTTCAG TTTGGGAACA GGCTACAGCT
50751 TTGGATTTTT CAAGGCCTAG CATTCCCATC ATTCTAAATT TTACTTAGCT
50801 AATACAATAG TAGTTGCCAG AGCTGATGAC ATAGTATTTT GTCATGCTTG
50851 GCTCCGTTCA AGCATTTTAG TTTTTTAGCC ATTACCATGG CTAGACCCAG
50901 TCAAAAGAAT TTTCATTGTT TAAGATTCCC ATTATCCTAG TTTTTACTAG
50951 TAGCCAGCCA AAGAAAAGAA AAAGGAGGTC AGAATTTCGG TATTTACATA
51001 GAAATTTAAG GGGAAAAGGC CAGGCATGTT TTTAAAGTGT GGAAATTAAG
51051 AACTATTCAT TATCCCACTG ATTGTGTGGA TGTGTTTTTT AAAGTTTTGT
51101 TACTGTCTTG AGAGAGAGAA TATTGAGATA GGACATAATG TTGGTTTAAG
51151 GGAATGAGGG TACTTTCTGT AGGTGAGGTG CCAAGCCATG TCATCAGAAA
51201 TGTTAGTCAC ATGACTTTCT AAGCACACCT TAAATGTTTT ACCGTGTATG
51251 TTTTTGTAAA GTTTTAAATT TTTAACTGGG AAAAACAGAC CTGTATATTA
51301 AGTTTTATAT ATATATATAA ATTTAAAATT ACATATATAT GTTTATATAT
51351 GTAACTTTTA TATGGGAGAG ATATATATTT CTATATCCTC TATAAAAAAA
51401 CATATCTATA TATGAAAATT ATGTACGTAA ATGTTAATTT ATAATTAATT
51451 ATATAAATAT TAACATAATT ACATTATATA TATAGAAAAC CTAGTGTACA
51501 GATCTGTATA TAAATTAAAA ATGTATGTGT TATATATAGT TACATCATAT
51551 AATACATATA ATTGATATAT ATAATGATAA ATACTTTATT GAAGGATGAA
51601 AAAATTTCCA TGCTGTCTCA TAAAATAAGA TGGTTGACAT ATGCTAAACT
51651 AGATAGATTC TCCTGTTTCA TACTAAAGCA GAATGTTGTA AAATATTAAA
51701 TCCAAATGAG ATGTCTCAGA TTAAGGCCAT TTCAACAGGA ATGCTGAGAC
51751 TTTAAAAAAA AAAAAGTCT GAGGCTGGGC GTGGTGGCTC ATGCCTGTAA
51801 TCCCAGCACT TTGGGAAGCT GAAGCAGGTG GATCACTTGA GGCCAGGAGT
51851 TTGAGACCAG CCTGGCCAAT GTGGTGAAAT CCCGCCTCTA CTAAAATACA
51901 AAAAAAATAC ATGGGTGTGG TGACGCATGC CTATAATTCC AGCTACTTGG
51951 GAGGCTGAGG CAGGAGAATC ACTTGAACCT GGGAGGTGGA GATTGCAGTA
52001 AGCCCCACCA CTGCACTCCA GCCTGGGCGA AGAGCAAAAC CCTGTCTCAA
52051 AAAAAAAAAA AGCCTGAATT ATATCAGCAA ATGAAAACTG TAATGTTGTT
52101 CTCTGTTTCA GAGGCCCTTG AATGAATAGC ACTAAAAATA TTTTTTAAAA
52151 AATGAAGAAA ATGAAAATTG TAATGTTCCT TATTTAAAAG GCCCTTGAAT
52201 GAGTAGCATC AAAAATATTT TTAAATGGGA GGCCAGGGTG GGAGGTTTGT
52251 TTGGCACCAG GAGATCAAGA CCAGCTTGGG TAACATAGCA AGACCTTTGT
52301 CTCTACCAAA AAAAAAAAAT TGGGTGTGGT GGTGCCACCT GTATTCCTAG
52351 CTACTGGGAA CACTGATGCA GGAGGATCCC TGGGACTCTA GAGTCCAGAG
52401 TGAGACCCTG TCTCTAAAAC AAACAAACAA ACAAAAACTG TATTTATGTA
52451 AAAGTAATAC TTGTTTTTTA AATTTTATTT ATTTTTAATT GATAAAAATT
52501 GTATGTATGT TTATGTGATG TATATATTGT GGAATGGTTA AATCAGGCTA
52551 ATTAACTCAG ATTTTTTGTG TGTGTGGGGA GAATATCTAA AATCCCTCTC
52601 CTTAGCAGTT TCCAAATGAA ATGAAAGAAT AAAAGTGATT TATTTTTTTG
52651 AGACAGCATC TCACCCTGTT TCTCAGGCTG GAATGCAGTG GCACGATCTT
52701 GGCTTACTTG ATCCTCGACT TCCCTGGCAT CCGGTGATCC TCCCACTTCA
52751 CTCTCCTAAT TAGCTAGGAC TACAGGCATG CGCCACCATG ACTGGCTAAT
52801 TTTTGTATTT CTTGTATAGG CAAGGTTTTG CCATGTTGCC CAGGCTGGTT
52851 TCAAGCTCCT GGGCTCAAAC GATCCACCTG CCTCAGCCTC CTGAAGTGCT
52901 GGGATTACAA GTGTGAGCCA CCACACCTGG CGAAAGTGT TATTTTTTTA
52951 AATGACAAAT TTAAGTCAAA GAGATTGAAT GTTCACTTCT GGTACTTTGT
53001 ATATAAGAGA AACATTCCAT TAAATAATTT TTTAAACATT TCTAAAATTA
53051 CATATTTTGT CATTAAATGT TTAAACAATC AGTATAATTT CATTGATACA
53101 GTGTTTGTTA TTTTGTCGGT GTTTAAGATT GATAATTGGG GTTAGTTTTA
53151 ATTCAGAATG TTATTCTATT TAATGTCACA CTTCATGTCT TTTTATTTTG
53201 TATATCTATT AATGAATTAT TTTAGCTATA GTTATTACTG TTTTAGAGAT
53251 GAGGTCTTCT ATGTTGCCCA GGGTAGACTT GAACTCCTGG GCTTCAGCAA
53301 TCCCCTCCTC AACCTCCGGA GCACATGAGA TTAGAGACGT GTGCCACTGT
53351 ATCTGGCCTG CTGTAGTTAT TTTTAATTCT TTTGTCTTTC AACTTTTATA
53401 CTAGAGTTAG AAATGATTTA CAAACCCTAT TGCAGTTTTA GAGCGTTATG
53451 AATTTGACTA TATATTTCTT ATAACAACTT AACTTCAGTT GCTTACAAAA
53501 ACTACAGAGT TTTACTCCCC CGTCCACATT TTATACTATT GATGTCACAC
```

FIGURE 3, page 17 of 27

```
53551 TTTACATCTT TTTATTTTGT GAATCCATTA ATGATACTTC TGGTAGTTTT
53601 TACACTCCAC TATTCAGTTG TCAGACACCA TTCAGTTGTT AGATTGTTAT
53651 GAGCTAAAAG CAACTTAATG GGTATTTTTC AAAAATCATT TATGTCAATT
53701 GCTAATGGAC TTCTTTTCTA TGCCATGATC ATGCTTTTTT TATTTTTGAG
53751 ACGGAGTTTC ACTCTTGTTG CCTGGGCTGG AGTGCAATGG CGCGGCCTCA
53801 GCTCACTGCA ACCTCCGCCT CCTGGGTTCA AGCGATTCTC CTGCCTCAGC
53851 TGGGATTACA GGCATGTGCC ACCGTGCCGG CTAATTTTGT ATTTTTAGTA
53901 GAGACAGGGT TTCACCATGT TGGCCAGGCT GGTCTCGAAC TCCTGACCTC
53951 AGTTGATCTG CCCACCTTGG CCTCCCAAAG TGCTGGGATT ACAGACGTGA
54001 GCCACTGCGC CTGGCCTGAT CATGCTTTTA AGGTGGTTGA GTAAGTACTA
54051 GTTGCTGGGG CTTTACTTAG TGCCCTCCTA CTCAAATGTG TTAGAACATA
54101 GTTAAGAAGG CTGTAGTGTT CAAAAGGAGT AAAAAGCAGT GCAGTGTTTG
54151 CAGTAATATC TGCTTCTCAA TTTAGGACTG ATGCTTATTA TGGCTTAAAT
54201 GTTTTTGTAG TAAAATTTGT ATTCAAAAAA TATATTTTTT TTTCTTTTTG
54251 CGACAGAGTC TTGCTTTGTC ACCCAGGCTG GAGTGTGGTG GTATGATCAT
54301 GGCTGACTGC AGCCCTGACC TTCCGGGCTC AAGTGATCTT TCCACCTCAG
54351 CCTCCCAATT ACTTGGGACC ACCAGCATGC TTGGCCGATT TTTTTTTTTT
54401 TTTTTTTTTT GTAGAAGCAA GGTTTCCCTA TGTTGCCAAG GCTGGTCTTG
54451 AACTTTAGGG CTCATGTGAT ACTCCTGCCT CGGCCTCCCA AAGTGTTAGG
54501 ATTACAAGCC TGAGCCACCA TGGCCGGCCA AAATATTTTC ACTATAACAA
54551 ATATCATATC TGTATATACT CAGTTTTAAT ACTAACTCAA AGTAGAAACA
54601 TAAAGCTGAA TGACTATTTT ATTTTCAGAT TCTCTCCATT GAGTTTCCTT
54651 CTCCGTCTTG TGTGATCTCT GAACTTTTCT CCATCTTTGC CACTTCTTGT
54701 CTAGCATTTT TTTTTTATCA GCAGTTTCAT TCAGATTTTT TTTTTAGTTC
54751 TTTCAACGGT GGAGTGGAAG TAGGCAGCAG GACAGAAGAA CTTGAAGCAG
54801 AGCACACTGG AGAGGAGAAA TTAACAAAGC CTTTATGAAT AAAACAACCC
54851 CCCAATATCA GTCTGTGTGC ATTATGAGCA TAATTGTACT TTCATCTCAT
54901 CTGTAATGTT CATGACTTTT CTAGAAAATT ATACTTTAAC ATGAGAAAAG
54951 AAAAAGAACC AGCTAATTCA TAGGGATGGA GGACACAGCA TAGTCAAAGC
55001 AAGAATGAAA CTCTCTTTAG TGCCACCTCC AGTGCAGAAT AAGTAACATT
55051 CAGCAGAGGC AGGTTTCATT TGATAATGGA TTCCTATAAT AAACTGCGCT
55101 CAGAATTTGT GCAGGTTTTA AAATCCCGTA TTCCAAACCC ACTTCCTTAG
55151 CCCCCAAGTT AGAAAACAGC TTCAGTAAAG AAAATTGTAC GATGATATAA
55201 CTTTACCAAA AAATAATTTC TTTCCATGAA GATGATATAT TATTGTTGAC
55251 TTCTAATTCA ATCAAATATA AACAATTGCT AAATGGCTTT TCAGTTGACT
55301 CCTTTCTTGG TTAAGGAGAA GATAGGAAAA AATGAAGGGA TCAGAAGTCA
55351 TAGGATACAT TAATTTTTTT TATCTCTGAA TAAACAGGTT GCCTACTTAA
55401 AAATCTATCA GTTTAAAAGT GTTGGTCTCT TCTCTCTCTT TTCAGAAAGT
55451 GCCAGGAAAA AACAAGAAGG TATCGTTGGC AGTTCTCGGG TGTGTTTTGC
55501 ACAGCACACC CCATCTCTTC CAGCAGAAAG TCCTCGGCCA TTGAAGCTTC
55551 GAAGCATTCT TGCATGAGC CCTTTTACAG TGACAGACCA CACCCCAATG
55601 GAGATCGTGG TGGATATTTT CCGAAAGCTG GGACTGAGGC AGTGCCTTGT
55651 AACTCACAAT GGGTAAGTCT GGTACCACAG GAATCAGTTC ACTTGCTAGA
55701 ATATAGGATC CTTTTTAGTG GAATCTATAT AGTTATTAGG GGAGCATGTG
55751 AGTCAGCTCC CAGGTGGGAA AGTCTGTCCT ATGGTATAGT CACAAATATA
55801 GGATCAGTCA ATCAAATTTC ACATTTACTA AGGAATAAGA AAGATGTCAT
55851 CTGCCTGCTC TTTGCCAAAC AGTGACATTT GTAAATAATA CCTCAAAGTT
55901 GGAAAAGAGG TGCTGAAAGA TCTCCAGCAT GAAAGCATGT TGAGCTTAGA
55951 GTGCTTCTTT TCCTAGGGAA GAGTGGACCT AACCTGCATG GAGCACTGCA
56001 AAAACCTGTT TTATTTTTGT AAATGTTTCA TTTTTAGTAT ATAAATTTCT
56051 AGTACAATAA TAAGTTTCTA GATATTTTGC TATTTACTCT TTCAGCCAAT
56101 ATTTGATTTA TCATGTAATG AAGGAAAGAA TATATACTTA AATGAAATTT
56151 GTAAATGAGC TAAAAATCTC CTTTAACAAA TGCTTTGTTT CCTTTTGTCT
56201 ACCTTTCTCT ATACACAAAT CTTTTATATT TATATAACTG CTAAGGACAA
56251 ATAAATACTC ATGTATTTAA AATGTATACA TTGATAATTT ATTTTTCCAC
56301 CTTTTACACA TGAACTGCCA GTGTTTCTCC ATTGACAGGA ATATAGGAAA
56351 GAAACAGATG TCACGGGGGT TGTGGAGACC TTAATGCACA GAATTGATTT
56401 AGCAAATACA CTACTTCGTC ACCACTGCTC TCTTTTCCTG GACCTGGGAT
56451 CTGTTTCTCC ACACTTCTTT CTTTAGGACC CTTCATTTCC ACTATATATT
56501 CTTTCTTGTT GAACTTAAGA ATGTTGTTTT ATCCGAAGGC AAATACCAAA
56551 AAACAGAGGG TATTCTTGGA TTATGCATAA ACTGGATGGC TAATCCTGAA
56601 CAGCGTAAAG CTGGTTGAAA TTCTAAACAG AGAATCATAG CAGTTTTTTG
56651 TTGTTTTTTT TTTTTAACAT GTTGTAGAAA ACACATTGGT GACAGAATAC
```

```
56701 ATGACTCCTG TCCAGAGAAA GGAGAGAAAA AGAACAGAAA GGAAGGAAAT
56751 TTGTTTATTG AACACCTTCA TATTTTCTCA TTTAACTTTG CAGGACCTCT
56801 GCAAAGTAGG TAGTTATATC CCTACTTTAC AGATGTAGTA ATTAAAGCTC
56851 AGGAAGCTTT AATAATTTGC CCAAAGTCAT GTGGTGAACA AGTCATGGTT
56901 CAAGGAATCA GACTGTCTTT CCTACTTTAA AACCCAGCCT CTTGCTACTA
56951 TTTTGCACTG TAAGTGACTG ATAGAAATCC TCTTTCTTTG TGATTTCTTA
57001 AACTACTAAA ACATTTTCTT GGCCAATATA TTAGATTGAG TTAAGAATAG
57051 AAATATGAAA CTAGAGAATT AGATCTATGT TTAGTGTTTT TCACTGCGCT
57101 AATTAAAATA ACTCTTTAGG AATATGAAGT AAATCATTAA AGAGATAAAG
57151 CCCTTAAAGG CAGGGAGTTT AGAATTATTA AATTCTAATA ATTTAGATAC
57201 TGATTGGAGA AGAGATGTAT TCATAAGTTA TTATTGTTAC TATTTGTCTT
57251 TGTGTAATAT TGTTTGATTA AATGATGGCA CCGACTTCAT TAAGTTTAAA
57301 AACTCAGTAC TAGTTAAATG GGGCAACTTT TCATAAAGCT TTGCTAGTCC
57351 TTGAGCCCTT TTATTTGTTA AATGGCTCAA CTGGAACCTA AGCTGAGTTG
57401 TTACAAACTA TTATTTGCTT CAAGTTGTTT TCTGTTCCTG GCATGGCTTT
57451 TTCTTTTGTG TACTGACAAA TATAAATGTT ATTCTGTTGA GTTATGGTTA
57501 ACTATGAACA CAGAACTGTT AGGGATTAAT TTTCATATTT CAGTTTGTTG
57551 ATTAATTCCC AGGTATTTGG CAGCATAGAT ATTAGAAAGG AAAATATTTA
57601 AAAGAAAGTG TAAAAATAAC GAAGTGTATA GAGCGAGGGG TGGATAGCTA
57651 ATTAAAATTT TGTCTGGTCC TGCCTGTTCA TATGAAAAAA GGGGTTGGAC
57701 TTTCTTCTAA GGGAATATAT TAAATTGCTT TCATCATATT TTCCTTATTT
57751 CTGTCTGTCA AGGAAAATAA ATTGATACAT ATATGGGGAG AAAAGAGATC
57801 ATTTAGGGAA GTGGCTCATG GGACTTTTTG TTTTGTTTGA AGTGTATTAG
57851 GAAGTCGGGT GTTTTTTTTC TCACTTAAAT TATTTAAAAC CCAGAAAAGA
57901 AATGATATCT TCTGGTTTTT AAAGGAGACC ATGAAGTTCT GCATAGCTAT
57951 CATTGATGTG TAGTTCATAC TGCATTTTTA GAAGTGGAAA ATAGTTATTT
58001 GGAGGAAGAT AACAAATCTG GAACCTTAGG TGCAAGGAGA AAAAGAATAG
58051 ATGAAAGGGA AAGATGTTTG TAAATTATAA AAATTTCAAT TAGCTATTGG
58101 TTTTTGCAC TTTATATTTT AACTGCAGAA TTTTTCAAAA TCAGTTAATC
58151 TTGGTGGAAT TAGCAGGATG TTAATAGGAG TGACTCAGAA AAAAACATTT
58201 TGTGACTGTC TAAGTTTGGA AAGTATTGGA TTAAATACAA TTGAGGTTTC
58251 TTTACTATGG AACTCCTCAG AACTTATAAT ATGTTGATAT TCTTTGATTC
58301 CCAGATGAGG GGATGGGTAA TAGGATACAT GGTTTTCCAG ACTTGTTTGA
58351 AAATGCAACT ATTTTTGGGT TGCAGGGAAG GATATAGTAG AACTCATGGG
58401 AACTGGTGTT TCTTGGAACA TGCTTTGGAA ATGCTGGGTT ATGCCCTGTT
58451 AACTCTTACA TCATTAGTTT TTAGCCCAAA AGGAAACAGC AAATAATGTT
58501 TTATATGAGC CACATTTTGC GTTGATTTTC CTTCCACTCT GTAAAATTAC
58551 TAAAGCAGCA CTCTGACTTT ATTATGCTCA AATCGCTCTT CTCCATTAAT
58601 GTGTGTTTCT CCATCTTTTA GGGTTTTTAC TTTATAAATA CAGAGATTAC
58651 TGTGTAAAAT TCTAAATTTG CCACTGGGTC GTTATACATT TGTAACCTTC
58701 CTCACAGTAT ATTTTGTGAT TTGGCAGAGT TTACCAATAT AGATGATACT
58751 AACTGAAATT AATCATTCTG TATAATTGGA TAGAAAGCA TGAGTAAGAA
58801 TTCAATTGGT ATTATATTTA ATTAATTGCC AAGATTTTCA CATTTCCTGA
58851 CTACAACAAT AAAATCAAAT GAATTGATGG CTTAAAAAAA AGAAATCTCA
58901 AATGTTTAGT CAATGAAGAA CATCTATTGA ATGAGTGAAT GTTCATTATA
58951 TATAGTGCAT TTTCTGAGCT TTTTTGGAGG GGGAAGTTGC TCCCATGCTC
59001 TGAGAACTTT TAAGGATCGA TACATTATTT TTAACATAAT AATGAGAAAA
59051 CATGAGCAGA GAACCCATTT CTGTCATTCC CATTCTCTAT CCTCCTGCTC
59101 CCCCACCTCC CACCCCAGCC ATCAAGCTAA GTAACTATTT TACACCTGGA
59151 CGTAGCTATA GGAACAGGCT ACTTTGAAGT CTCCTAGTGA CATCCTTCAA
59201 GTCTGAATGT TCAAAGGCAG TTTAACAGGG AGGTTGACTT AATGAGATCA
59251 TCAAGGAAAT GTCCAGTCAT CCTGAAGGGT ATTTTGGATG GGCTTCCAGA
59301 ATTTAAAGAT TAAAGTTTTT TTAAGGTTTT TTTATTTTCA CTGTTTATAT
59351 TGCCACATTA ATTTCCATTA TAAAACCAGT AACCATAGTT TTGTTTTAAT
59401 TAGCAATCTA ATTATTTTCA TGTATCCTCA TTATGAGAAT TTATGTCCAT
59451 CACTTTGCTT GATGTGATAA CAGTGCATG CTAAATGAGA AACAATTGTT
59501 ATTTAGAAAA AAATGCACAA AGTGAAAGTC CTTTTAATCC CTAATCATAA
59551 ATACATTTTA TTAGCTTACT TTAAGAAGTG GCAGTCACAG CTCCTGAACA
59601 TTAGGGAGTG TTTCTTTTGG TCAGCATTAT TTATTTAGTG CACATTGCCT
59651 TTAATTTTAA TTTGAAATTA TAGTAAAATC CACGGGAGTT TTTAAGTCTC
59701 CTCACAGCCT TTTGCTACCT TTTCACCAAG GTAGATCCAG ATGATAACTG
59751 CTGTGTTGTG ACATCATAGA AATTAGAAAA ATATTTTCCT CTGAGGAAAG
59801 AACATTGTAA ATGAAACTCT ACATATCAGA GGTCTATAGC TATGTATCAA
```

```
59851 TATTAAGTTT CTTTTGTACT TTGCTTTGTA GTCATCTTCA TTCCAAACTT
59901 TCATAATTAT TATTTTTACT TTAAAAAGAA AAATAACCCA CCAATATTGA
59951 AGATTAGTAT TGTGTCACTT TTGAAAGTCA GTAGAATTTA TGCAAAAGGA
60001 ACCTGGAACT TTAAATCATT TTGTTTTTAT TTTCTAAAGT TCATGAGACT
60051 CATTCTTATG GTTCATGTTT TTATTTTTTC TCTCATTCTT TATCATTATG
60101 ATTGGAAACT CTTTTAATTT AATTTCTCAC ACAGTTATTA GCATAATAAT
60151 CTGTTTCAGG ATTGTCTTGG GGATCATCAC AAAGAAGAAC ATATTAGAGC
60201 ATCTCGAGCA ACTAAAGCAG CACGTCGAAC CCTTGGTGAT TAGATATATC
60251 AGATCTCCTC ATTAGACACC TTAGAAGTCA GGAAGCATGA AACTTGTGAA
60301 CTGTTGAGTT CTGTCTTTCC CAGATATCTG CTGAACAAAA ATATCCTACT
60351 ATGCTGCCAA TTACATTTGT ATCTGATAAA ATGTGTCTGT AAGATAAATT
60401 TAGATATGTG TAAAATCCCA TTTATAGAAA GTAAGCAAAA GTTAACATCT
60451 CTCATCAAAT CATTCATTAC AATTTCAGAA CTGTAAACAG TTTGGTAGTG
60501 GAATAAGTGA ATATTATTGG ACATTCTTAA AGTGAATATG GCAAATCTGT
60551 CTACCTCAGT GGATACACCG GTCTCAGAAG ACACCTGACT GGTTAAAAAT
60601 GTCTGACCCA TCCCCGCAAG CCCTTTTTTT TTTTTTTAAA TGTTTCCCGA
60651 TCTTGTGGTA GTCTTATGGT AAATCTAAGC TCCTAAAGGA TTTTAAAGGA
60701 GCTTAGCAAT TAGAACTGCT TACAGTTAAA TGGATTTTTT AATGGGCACA
60751 CTAACTAGAG TGTAATGTGT ATATTATTTG TGATCATAGC ATTAGTTCTT
60801 TTTCTGCTAT ACCCTGCATA TCTTCAAAGT CACAGTGTGT GTCCTGCCAT
60851 CTCATTAGTG AATTGTACCT AGATTATGTG TGTGCCCCTT TTGTATGATG
60901 TTTCTGGAAC GCTATAAGCA GCTTTTAGAG TCAAATGCAT TCATTTTAAC
60951 TGGCTTTATG TCCTAGTGGT TTCATGACTA CAAATTTGAA TTATCTTACT
61001 GCATAACATA AAAAATGTCT GGCTTTAGCA ATTAATGCCC GAAATTATTT
61051 TGCCCTGCAA TTGTCATACC TGTATGAAAC CTGTCCCAGT TTGCTTAAGT
61101 GCACAACTGA TTATGTATTC CTGTGTGTAT GCTAATATTT CACAAGTGTT
61151 TCATGCATCC TTTTTTAAAA AACTACTAAC CAGAATATTA TCGTAGCTAC
61201 TCATTCATTC TGCTTTCTGC TTCACCTATA ATAATCTTTT AGGACTGCCT
61251 TCTGATTTTT CACCTATCTT TTAATGTAAG CATTAACAAC TAAGACTTTC
61301 ATAAAAGCAC TGTATCTTAA CTTTCCTGGC CTAAATCAAA AAAAGGAAAA
61351 CATTGATAAG TGTCCTAGAA ACTTGGATTC TTTTATAGAT TTGTTCTTGG
61401 GGCTCTGATG TTTGGGATTG ACGTTCTGTG CTGACCATTT TATATGCATT
61451 TTATCTTAAT AGTATGTGCT TTCATGAAGA TTCTGATACA AGTGGGCAAT
61501 CCTTAAATTA TCTTTGAAAA ATTGGTTAAT TTTGGTTAAA AAAGGGAAAG
61551 TGGCTGGGTG CAGTGGCTCA CGCCTGTAAT CCCCAGCACT TTGGGAGGCC
61601 GGGACGGGTG GATCACAAGG TCAGGAGTTG AAGCCCATTC TGGCCAACAT
61651 GGTGAAACCC TGTCTCTACT GAAAATAATT GGGGCATGGT GGCACATGCC
61701 TGTAATCCCA GCTACTTGGG AAGCTGAGGC AGGAGAATTG CTTGAACCGG
61751 GGACCCAGGA GGCGGAGGTT GCAGTGAGCT GAGATCGCGC CACTGCACTC
61801 CAGCCTGGGC TACAGAGCGA GACTCTGTCT CAAAAAATAA ATAAATAAAT
61851 AAATGAAAAA GAGAAAATAT TGAGAGGATT TGGTCATCAT TTTACTGCTC
61901 TCTTCATGTG ATGGAAATCA ATTTTCCTTC TCAAATGGGA TCAGTATCAT
61951 TTCCTAGTCA TACATCCATC CAGTTTTTGT TACTTTTTTG TTGGCATACA
62001 TTAATCAAAA TAGCTCTGCT TCATTGAGGC ATGCAGTCCT CAGACTCTCG
62051 GTGGAAAGGC TGTCATACTA TTAGTGACCA TAGTAACTTT TTATACCAAA
62101 GGATGGTTGC TGGATAATTT TAATATCTTT ACCAATAAAG TACTTTTTGG
62151 AAATACAAAA TCAGGCTGCT TGCTTTGCTC TATTCCTGTC AACAAAAAGG
62201 ATTTAGCTAT AGATTTAGCT TCTCCTTTTA TTTTCCCTTT TATTTCATAG
62251 GAGTCTTCTG TTTATTCCTT TCAGGCGCCT CCTTGGCATT ATAACAAAAA
62301 AAGATATCCT CCGGCATATG GCCCAGACGG CAAACCAAGA CCCCGCTTCA
62351 ATAATGTTCA ACTGAATCTC ACAGATGAGG AGAGAGAAGA AACGGAAGAG
62401 GAAGTTTATT TGTTGAATAG CACAACTCTT TAACCTGAGG GAGTCATCTA
62451 CTTTTTTTTC CTCCTTTACA AAAAAGAAA GGAAATATAA AAGCCGGGTT
62501 TTTGCAACAT GGTTTGCAAA TAATGCTGGT GGAATGGAGG AGTTGTTTGG
62551 GGAGGGAAAG GAGAGAGAAG GAAAGGAGTG AGGTATTTCC CGTCTAACAG
62601 AAAGCAGCGT ATCAACTCCT ATTGTTCTGC ACTGGATGCA TTCAGCTGAG
62651 GATGTGCCTG ATAGTGCAGG CTTGCGCCTC AACAGAGATG ACAGCAGAGT
62701 CCTCGAGCAC CTGGCCTGTT GCTCCAACAT TGCAAAGACA CATTATCAGT
62751 CCCTATTTCT AGAGGGATTA CTTTGAATTG AGCCATCTAT AAAACTGCAA
62801 GGTCTTGCCC TTTTTTTTAA TCAAAACTGT TCTGTTTAAT TCATGAATTG
62851 TATAGTTAAG CATTACCTTT CTACATTCCA GAAGAGCCTT TATTTCTCTC
62901 TCTCTCTCTC TCTCTCTCTC TCTCTACT GAGCTGTAAC AAAGCCTCTT
62951 TAAATCGGTG TATCCTTTTG AAGCAGTCCT TTCTCATATT GAGATGTACT
```

FIGURE 3, page 20 of 27

```
63001 GTGATTTTAC TGAGGTTTCA TCACAAGAAG GGAGTGTTTC TTGTGCCATT
63051 AACCATGTAG TTTGTACCAT CACTAAATGC TTGGAACAGT ACACATGCAC
63101 CACAACAAAG GCTCATCAAA CAGGTAAAGT CTCGAAGGAA GCGAGAACGA
63151 AATCTCTCAT TGTGTGCCGT GTGGCTCAAA ACCGAAAACA ATGAAGCTTG
63201 GTTTTAAAGG ATAAAGTTTT CTTTTTTGTT TTCCTCTCAG ACTTTATGGA
63251 TAATGTGACC GGGTCTTATG CAAATTTTCT ATTTCTAAAA CTACTACTAT
63301 GATATACAAG TGCTGTTGAG CATAATTAAA TAAAATGCTG CTGCTTTGAC
63351 AGTAAAGAGA AGGAAGTATT CTGATTAGCT GTATCTGGTA TTAATTGCAT
63401 GTTAAAACAC TGGAATTTTT AAAATTGAAA TTAGATCAGT CATTCTTTTC
63451 TTTTCTCAAG ATATCTCATG GCTGACACTG AAGAAGAAAT GTAATTCATA
63501 ACTTGCACTA AATGTATATT TTTTTTCTTA AAAATTTACC ATTCTTATTT
63551 ATATTTTTAT GGATTAAAAT TTATAAAATA CAGATCAGTT AATATTGCAC
63601 TTAAGTAATT TTACCTTTTT AATGTGATTT TTATAGAATA ATTCAGACTT
63651 ACAAATACAG AGATATGAAC AAAGTTTACA GTGGGAACAA AGGTTTAAAA
63701 AAAGGTTGTG GTTCTCTCTC TGTGATCCAG TGTGCACATA AACCTTTCTC
63751 TGATCTTTCA CTGCCATCCT CTGGATTATG TCTTCTGACC TGTCCATTTT
63801 GACCCATTAA CTGGAAAGTT GAAAAACTAC ATTAACTGGA AAGTTGAAAA
63851 ACTACATTAC TTTGGAGAAT AAAACCGAAA GTTCGTGTAT ACCTTCTTAA
63901 AAAAAAAATC AAACCAAAAA TGTGAAAACA ATAGAATTGC AAAGATAGCA
63951 GTTAAAATTT TAATCTGAAA ATAACCTTTG AATCTCGGGC TAGGTTATGT
64001 CCATATTTGA AGTGGTCAGT GATGGTTTGA ACATTTTTTG CAGGATGAGT
64051 TAAAATGCAC TGGATTATAT TTGGGATTTT TGTTTTTGGA ATTGTCTGTT
64101 TTAATCACAG CCTTAATTCA CAATTGGCAA AGGCAGTTTA CTCAAAGGAC
64151 TGGGCTAAAT ATTCTGTAAT TATGCATTTT TGATAGGAAA ATGAAATTTT
64201 TGCAAACAGA CATTTTCTTT TTTTTTGGCT GGAGTGCAGT GGGGCATGGT
64251 CTTGGCTCAC TGCAGCGTTG ACCACCTGGG CTCAAGTGAT ACTCCCGCCT
64301 CAGCCACCCA AGTAGCTGGC ACTACGGGCA CACGCCACCA TGCCCAGCTA
64351 ATTTTTTTGT ATTTTTAGTA GAGATGGGGT TTTGCCATGC TGCCCAGGCT
64401 GGTCTCAACT CCTCAGCTCA AGCAATCTGC CTGCGTGAGC CTCCCAAAGT
64451 GGTGGAATTA CAGGCGTGGG CCACTGCGCC TGGCCCAGAC AGACATTTTC
64501 TGAAACACAA CTGGCAATGA GCTGTTTTTA CATTTTGAAA GTGATTCTTC
64551 ACTTCCTAGT TCTTAATTAT AGTATACCTA TTAAGATCTG TAAGATCCTG
64601 AAGACATAAG ATCATGAAGC CATATAAGAA TGAGGATTGA AAGTTGAGCA
64651 AAATTTTCGG GATTTTGGGA AACATTCTTA GCTGTGCTAT CTGCCTAAAA
64701 TTATTCCTTA TTACTTCTCT CCTTTGACAG ACTTCAAGTT TTCTTCATAG
64751 CCCTTTCAAA GTTTTTTGAG CCATCCAGAG TAAAATCATT TCTAAATGAT
64801 AGTTCTGTAT ATCTCCAACT CGTCTTAAGT GTATTTGCCT GTGTGCAACG
64851 TATTGCTAGA CTATGAACTC CTCAGCATGG CTGCTGGATA ACTTAATTGT
64901 CCTGAGTTAA TAGCCTTCAA AGGACAAATC GGTTTCTTTG CAGATAGCTT
64951 CGTAAAACTT CACATGGAGT TTATTTTATC ATATTTCCCT TTTTTATTTC
65001 TGCTCCTCCT TTAATTGCCC ATCTTGCTTC AGAGACTGAC ATTTCAGGGT
65051 GGATATTAAT TAAAGCATTA ATTTTGTTTT TTGGTATATT TCTATCCCTA
65101 GTATTTCTAT CTTACTGCTA AAATACAGGA AAAGTGCCGT ATTTTTAATG
65151 CATTTAGTGG TTTTCTTTGG TGTTATCTGT TCCATTTTTC TTTTTCATAC
65201 ATTGAAGTGT GTCTCCTTTT CAACCAAAAT AATGAAATAG TGGAGACCAT
65251 GAAATTGTTG TGCCTGGCTA ATTGGCAAAT TAATTTACCA ATATAATAAG
65301 TGTAGCGCCT TGTTTGAATA CCCTTTTTGA GAAGGTATGA TGAGAATGGG
65351 CAAGGGTGT   (SEQ ID NO:3)
```

FEATURES:
```
Start:    2159
Exon:     2159-2237
Intron:   2238-22041
Exon:     22042-22199
Intron:   22200-30359
Exon:     30360-30459
Intron:   30460-31475
Exon:     31476-31663
Intron:   31664-32964
Exon:     32965-33087
Intron:   33088-34548
Exon:     34549-34755
```

```
Intron:   34756-37975
Exon:     37976-38056
Intron:   38057-39552
Exon:     39553-40098
Intron:   40099-46366
Exon:     46367-46553
Intron:   46554-49237
Exon:     49238-49636
Intron:   49637-55445
Exon:     55446-55662
Intron:   55663-62274
Exon:     62275-62362
Stop:     62363
```

CHROMOSOME MAP POSITION:
Chromosome 4

ALLELIC VARIANTS (SNPs):

DNA Position	Major	Minor	Domain
1275	T	C	Beyond ORF(5')
1456	T	C	Beyond ORF(5')
5893	G	A	Intron
6226	A	G	Intron
8866	T	G	Intron
10397	C	T	Intron
10621	T	–	Intron
19651	A	G T	Intron
19891	T	–	Intron
20272	C	A	Intron
20412	T	A	Intron
23340	A	G	Intron
29948	T	A	Intron
33579	A	C	Intron
40762	G	A	Intron
40936	T	C	Intron
45998	A	G	Intron
47771	T	C	Intron
48117	C	T	Intron
54563	T	G	Intron
58735	C	T	Intron
59643	C	A	Intron
61638	G	T	Intron
63291	G	C	Beyond ORF(3')
63463	A	G	Beyond ORF(3')
63636	G	A	Beyond ORF(3')
63998	T	C	Beyond ORF(3')

Context:

DNA Position

```
1275   GCATTTCAGGAGGAGAATCTCCCAGTCTAGAGGAATCCTCTCAGAGGTAGCTATAAAATA
       TTGAACTCTGATCTTCAATAAGCATTGTGCGGTTTTTGTTTTTGTTTTTAATGACAGTTT
       TAAACAAGAAAGTTGCTTTATTTCTGAACTTCATAAAAATTTCTATTAAAGAGACAATTT
       CTGAATTTTATAACAATTTCTAGAACAGTTGAGTACCTCACTTTGAGACACATTTTTGCT
       AAAAGTTAAAAACACAAAACCCTTATGAGATAAAATAGGAAGCTAGTAGAGATAGGAAAG
       [T,C]
       CCTCTGCTTAGTAAACCTCTTTTTTTGCGTAGTTTAGACACATACAATAGTAAAGTTACTT
       AGTACGTTGATAGTTTTCTTTCTCCTCAAAAGCTACAATGTCTTACTAGCTAGTTCCTTC
       AAGAAAGGAAACAAGAAGCCGCTGGAGGAGATTGGTGAGTGGGATAAAACACTATTCAAC
```

<div align="center">

FIGURE 3, page 22 of 27

</div>

```
            TCTTCAGTTATTCGGTTTTTAAATCCTCAATGAAAGGCTGCTGTATTATAGAGTATTTTT
            TTTTTATTTTTAATAGACTTAGAACCAAGTTTCTTGAGAAACCTTTGGCATATTGTAGTT

   1456     TGAATTTTATAACAATTTCTAGAACAGTTGAGTACCTCACTTTGAGACACATTTTTGCTA
            AAAGTTAAAAACACAAAACCCTTATGAGATAAAATAGGAAGCTAGTAGAGATAGGAAAGT
            CCTCTGCTTAGTAAACCTCTTTTTTGCGTAGTTTAGACACATACAATAGTAAAGTTACTT
            AGTACGTTGATAGTTTTCTTTCTCCTCAAAAGCTACAATGTCTTACTAGCTAGTTCCTTC
            AAGAAAGGAAACAAGAAGCCGCTGGAGGAGATTGGTGAGTGGGATAAAACACTATTCAAC
            [T,C]
            CTTCAGTTATTCGGTTTTTAAATCCTCAATGAAAGGCTGCTGTATTATAGAGTATTTTT
            TTTTTATTTTTAATAGACTTAGAACCAAGTTTCTTGAGAAACCTTTGGCATATTGTAGTTT
            TTTTATGGCTATGACTCACATGACATTACTGTATAAAACTAGTACATTCTCTCGTAAAAC
            CACACAAACTTACTAGAGTGCTGCTCTCATTTTTCTACATTAGAAATGAAAAAGGGCATT
            GTCTGCATTCAAAATTTCCTTTTTACATCTCTGTATTACTTTTTCCCCTTTATATTTATC

   5893     TCTAGTTGACAAGACTGAGGTAAGGAATTGTTAAGGAAAAGTCAGAATTCCATCCAGATA
            TTTGGCTCATACTTTAATCATGAGGCTAAACTGCTTCTCTCTACACGTATCTTCATAGTA
            ACTTGTGTTTTAAGTCTGGTAGAAGCATAAGAAGTTTAAACACAGACAGAATCCTGTGGA
            AGTTAGTAAATTTCTAGTGAACGATAGAAATGATAGAAATCTCTTCTTCCCCCAAAGTCC
            CAAGAACAGATTAGTCTGCTTTTGACAAGTGTTATCAAAGTAGACTGTTCTCACATACAC
            [G,A]
            GGGGACTCAATAGGGCATTCCTGGTGGATATAATAAAATGAGTAAATGCGATAACAGGAG
            GAAATGCCTAGTGTGTTGCTCTTGGATTAGTTTTGATACAACAAAGGCAGCTTTGTTGTG
            AGTCAGTAGAGAGGGTAGTGTAGAAAGGTGGAAGTTGGAAGAGTGGCAGATCCTAGAGGA
            CTAATGATGGGCTTAAACCACAAAAAGTGTCGCTTTGCCATTGAA

   6226     ATAAAATGAGTAAATGCGATAACAGGAGGAAATGCCTAGTGTGTTGCTCTTGGATTAGTT
            TTGATACAACAAAGGCAGCTTTGTTGTGAGTCAGTAGAGAGGGTAGTGTAGAAAGGTGGA
            AGTTGGAAGAGTGGCAGATCCTAGAGGACTAATGATGGGCTTAAACCACAAAAAGTGTCG
            CTTTGCCATTGAAATAAAAGTTTGGGGTCTTATTTTTTCAATTTTCTCCCTGAAATTATT
            TCTTGACATTCATTAGCTCAGCAGTGTATCTAAATAAAGCTTTTTTGGGTTTCTATTATA
            [A,G]
            TAGAGGTTTGTTCCTTTTTCTTCCCTTTGAAAAGTATCATTTTTTGCACATTATTTGAAA
            ATCCAGGTGTTATATGATATTCTTATTGCCAGAGGGACATTCTGCAGGCTCTTTGTAAAA
            TGATTTTAGGATTCAGATACTTATTATATTTTTATTGGCCCTAATATTTTATCCAACTAG
            AAAATTAAACCTCTTCTTAAAAATTAATCCATCTAAGTGTCTGTAAATTAAAGGAACAAC
            TAAAGATTCTTTATTTGGTGTCAGAAACTCCTTGTTTCTACAACAGTAGTATAAAACAAA

   8866     ACATGTAAACCAACAATGAAATTATTTTAGTGACTTGAGAATCAAAGTGCTAGAGTTTGA
            ATCCCTGTTCTACTACTTGCTAGCGGTGTGACCTTGGGCCTGTTTAACTCTTGACACCTT
            GTTTTCCAAATTTATAAAGTGGAGATAATAATATCTGTCACATTGTGTTGTTGTGAGGAT
            TATATGAACTAATATATGTAATGTCCTGAGAACAATGTCTGGTACACATTAAGTTAATTA
            AAATTAGCTGTTCTTACTGTTATTATTAGACATGAGCTAGATAACAGTGGCCTCTACATG
            [T,G]
            GAAAGATTATTTTAATTCTGATGTAGTTCAGTTTATCTATTTTTTTTATTTTTGTCCCTT
            TTGCATTGATGTCATATCTAAAAAAACCTGCCTAACTCAGGATCACAAAAATTTACTCCTG
            TATTTTATAATTTTAGCTCTTTAGATCTAGGATCCATTTTTAGCTAATTTTTATATATGG
            TGTGAGGTAGGGGTACGGTTTCATTCTTTTGCACGTGAATAGCCAGTTGTCCCAGCATCA
            TTTATTCAAAAGACTATTCTTTCCTCACTAGAAAAAATATTTCTTTAAAGAATAATGAAT

  10397     CCAGGCTCCCTTGAACTCCTGGGCTCAGATGATATAGCCTCCTGCCACAGCGTCCTGATT
            AGCTGGGACTACAGGTGTGCACCACTACACGTGGCTTTCCTGATGAAATTTTAAATACCC
            AAATATTTGAGCAGAAATAATAGCTTGTGTTTATTGTTTTTCTACTATCTGTCAAGTATA
            GTATTAAATGTTTTACATAATTTGTCTCCAGTCCACATACAATACTCTAGTAGAAGTGGG
            TAACAAAACCAAGGTACTCAAAGAGGGTTAATAAGTAACTTGCGCTGGATCACAGAACTAA
            [C,T]
            GGGAGGCAGGGCTGGAATTTGACTCTAGGTCTTTCTGACCTCAAAGTGCAGTAAAGTCAT
            GGAATTTCTCTACTAGGCCACCTGGAAGAAAAGTGATCTTTTTTTCCAGTCTTTTTTGTTA
            CTGTTTTTCAGCCAGGAGATAGTAGAGTTAGGTAGTAGAATAGTAGTCACTGGCATCCGG
            TAGTCAGCCCTCCAAAAAAGTTTTTGATTTTTTTTTTTTTTTTTGTCTTAAACTTGGAAG
            CTACTAACTTTCAGGTCATACTTTCTTATCATCCAAGAGCTGGATATTTAGGTAGCAGAA

  10621     CTCTAGTAGAAGTGGGTAACAAAACCAAGGTACTCAAAGAGGTTAATAAGTAACTTGCGC
```

```
           TGGATCACAGAACTAACGGGAGGCAGGGCTGGAATTTGACTCTAGGTCTTTCTGACCTCA
           AAGTGCAGTAAAGTCATGGAATTTCTCTACTAGGCCACCTGGAAGAAAAGTGATCTTTTT
           TCCAGTCTTTTTTGTTACTGTTTTTCAGCCAGGAGATAGTAGAGTTAGGTAGTAGAATAG
           TAGTCACTGGCATCCGGTAGTCAGCCCTCCAAAAAAGTTTTTGATTTTTTTTTTTTTTTT
           [T,-]
           GTCTTAAACTTGGAAGCTACTAACTTTCAGGTCATACTTTCTTATCATCCAAGAGCTGGA
           TATTTAGGTAGCAGAAACTATGGAATTATCCTAAGTCCTCTTGAAGCTTCAGCTGTTAAA
           ATTAATTGGTTCTGATTAACACTGTGCTCAAGATTTACATTTCTAGGAGCCACAGTTTGA
           TTGGTCTAACTTGGATCTATGTGTTTTCTTTAGCTGGGGAGGAGAAGGTATCTTGATTGA
           TACCTTCACCAGGACTGCATGCAGTGAGGGACAGAAGTTTCCTTAAAATAATTGGGTTCT

     19651 TTTATTTTCTGCTACTATGGCAGAATTGAGTTGTTGCAACTGTGTGGCATCCAAAGCCTA
           AAATATTTACTCTCCTGGCTCTTTGCCAACCCGTTTTAGATTATGAGCACTTTGGCATTA
           TTATGTTTTTGTTTTCTTTCTATAGCACACAGTAAGATGTTCTGCCCACATTGTGCATAA
           TTTATGGGTTTATTCAAGGATTTATGCAAGTGTAGCTGCAAGAAAAAAACCTAGAAGTGA
           ACTTGCTAGGTTGAAGAGCA
           [A,G,T]
           CTGTGTATGTTAAATTTTGTTAGCTTTCGCCTTCCCAAAGGGATTATTCCATTTCATACT
           TAAACTACTAATTTTGTGATAGGACTTCTTTCTCCATAGCTTTGCTAAATTAATGCATTC
           ACACACTTCATCTTTACTAATCTGATAGAGGGAAATGATATTGTGGATTTGATTTGCATT
           TCTTTTTATGTGTTAGCTTGAGCTTATTTTCATATTTAAAAGCCAATTGTATTTCTTTTT
           CTTGAGCTATCTTTTAATGT

     19891 TTTATGCAAGTGTAGCTGCAAGAAAAAAACCTAGAAGTGAACTTGCTAGGTTGAAGAGCA
           TCTGTGTATGTTAAATTTTGTTAGCTTTCGCCTTCCCAAAGGGATTATTCCATTTCATAC
           TTAAACTACTAATTTTGTGATAGGACTTCTTTCTCCATAGCTTTGCTAAATTAATGCATT
           CACACACTTCATCTTTACTAATCTGATAGAGGGAAATGATATTGTGGATTTGATTTGCAT
           TTCTTTTTATGTGTTAGCTTGAGCTTATTTTCATATTTAAAAGCCAATTGTATTTCTTTT
           [T,-]
           CTTGAGCTATCTTTTAATGTCCTTCCTGATACATTTCTGAAGTCTGTGATACTCATATAA
           GATATATGGTGAACATGTGTCAAAGATTTATTTGACTCTAATGAGGGAACCCGCCTGATG
           ACAAGGCTGATTGAGAAGAGGATGTGTGAGATGAAGTGTATATCATCAGTGAAAGAAAGC
           AAATTCTTACAGGGCAAAAACAAAACCACAACTCTAAGGGTTATTGTTTCTACTGGACAG
           AATTCATTTGCATTTTACCAGATAAAAATTACTATTTTCAATTTATCTTTTACAAATCAT

     20272 CAAAGATTTATTTGACTCTAATGAGGGAACCCGCCTGATGACAAGGCTGATTGAGAAGAG
           GATGTGTGAGATGAAGTGTATATCATCAGTGAAAGAAAGCAAATTCTTACAGGGCAAAAA
           CAAAACCACAACTCTAAGGGTTATTGTTTCTACTGGACAGAATTCATTTGCATTTTACCA
           GATAAAAATTACTATTTTCAATTTATCTTTTACAAATCATTTTCTAATTTTACAGAGTCT
           ATTCCCTAATCAGTAGTAAATAGTCTTCAAAATTCTCCGCAGCGTCAGGTGACTATTATG
           [C,A]
           AGGCTAATTGTTGACACTCGGGCTTGACTTTAAGAGAACATGCCATAATCTTTTGGCCTT
           ACTTCCAAGTTTTGGATAATTTTTCTTAACACATTTTTCTCTAATTGCAATGATTTCAAG
           TGATATTATTTCTTTTTTTTTAAATTTTTTTACTATTTATTGATCACTCTTGGGTGTTTCT
           CGGAGAGGGGGATTTGGCAGGGTCATAGGACAATAGTGGAGGGAAGGTCAGCAGATAAAC
           ATGTGAACAAAGGTCTCTGGTTTTCCTAGGCAGAGGACCCTGCGGCCTTCCACAGTGTTT

     20412 TTATTGTTTCTACTGGACAGAATTCATTTGCATTTTACCAGATAAAAATTACTATTTTCA
           ATTTATCTTTTACAAATCATTTTCTAATTTTACAGAGTCTATTCCCTAATCAGTAGTAAA
           TAGTCTTCAAAATTCTCCGCAGCGTCAGGTGACTATTATGCAGGCTAATTGTTGACACTC
           GGGCTTGACTTTAAGAGAACATGCCATAATCTTTTGGCCTTACTTCCAAGTTTTGGATAA
           TTTTTCTTAACACATTTTTCTCTAATTGCAATGATTTCAAGTGATATTATTTCTTTTTTT
           [T,A]
           AAATTTTTTTACTATTTATTGATCACTCTTGGGTGTTTCTCGGAGAGGGGGATTTGGCAG
           GGTCATAGGACAATAGTGGAGGGAAGGTCAGCAGATAAACATGTGAACAAAGGTCTCTGG
           TTTTCCTAGGCAGAGGACCCTGCGGCCTTCCACAGTGTTTGTGTCCCTGGGTACTTGAGA
           TTAGGGAGTGGTGATGACTCTTAATGAGCATGCTGCCTTCAAGCATCTGTTTAACAAAGC
           ACATCTTGCACCGCCCTTAATCCCTTTAACCCTGAGTTGACATAGCACATGTTTCAGAGA

     23340 TTTTTTTTTTTGGAGGTCGGGGGACTGTCGCCCATTCTGTTGCCCAAACTGGAGTGCAGTG
           GTGCAATCTTGGCTCACTGCAACCTCTGCCTCCCAGGTTCAAGCGATTCTTGTACTCAGC
           CTCCTGAGTAGCTGGAATTATAGGTGTGTGCCATCATGCCAAGCTAATTTTTGTATTTTT
           AGTAGAGATGAAGTTTCGCCATGTTGGCGAGGCTAGTCTCAGACTCCTGGCCTCAAGTGA
```

<div align="center">

FIGURE 3, page 24 of 27

</div>

```
        TTGGCTGACCTCAGCCTCCCAAAGTAGAAAATCTTCTTGAAAAATAAAATTCCAAATCTC
        [A,G]
        AAAGGCCCTATATAATTTTGGTGTTGGAAATTTACTTGTCAATGAAAATGACTATTTACA
        CAAATTATAAGCTTCCATATTAATATATATGTGTGTGAACCTGAAATTCAAATTTTATTA
        TATTGTTTATGAAAGGTACAGCCTCTGAGATTCATCAGATGGTATTTACCTTTAGGGCAT
        ATCTAAAAATAAAATACAGTACATGAAATCCAGTGCTTTAATCCAGTGATTCTTAAACTT
        TTTGCTCTCAGATCCCCTTTAAACTCTTAAAAGATATTGAAGAGCTCCAAGGAGGCTTTG


29948   GACTCTACCAATGGGATCGGAGCTCTCCAAACCTGCATATTAAAAGGCCTATAAGTTTTG
        GGGGGTCCCTTTGTCCACATGATTATTCTGTAATACATTGTATTTATGGACATGGTATTA
        TTATACACAGATCCTGTCTTTTAAAGAACATTATAATCCACTTAACTGCTAGGACCAGAG
        AATGACCGATAATTCAAACCATATTGTCTTACAGAAGACATATATAAAAGATGGTTATGT
        GTACCAATTGAGGTTCAAATTTGATTCAATTTAAAACAATCTAGGCCAGATTTTATATAG
        [T,A]
        TTGTGGACCCTTTGCACTCAAATCTCAAGGTTCTTATTAAAATGCAGATCTTGGCTGGGC
        ACGGTGGCTCACACCTGTAATCCCAGCACTTTGGGAGCCCAAGGCAGGTAGATCATTTGA
        GCTCAGAAGTTCAAGACCAGTCTGGCCAACATAGCGAGGCCCAGTCTCATTGAAAGAAAA
        AAAATTTTTTAATAAAAAATAAAAGCAGATCTTGGGTAAAGACATGTAGTCTGGTTTACA
        GGTATTAACAACTGTCTGTAATGTAGTGATTTTGCTCCAGACTTACCTTTTCCATTATTT


33579   TTACTGTGAAGGCTGATTTTTTTTTTCTCTCACCACTAATTTAACACATGACTAGGCAAA
        TTTTCAGACTATTTAGTTAAACATCAAGAGCCTGGAAGAAGTATCTTGTGACCTAATGTT
        CTTTGACGGGTTAGTTGTTACTTTGCTGTTATGACCCTGAATTTTTTTTTTTTGAGACTG
        AGTCTTGTGCTGTCGCCCAGACTGGAGTGCAGTGGCGCAATCTCAGCTCACTGCAACCTC
        TGCGTCCCAGGCTCAAGCAATTCTTGTGTCTCAGCCTCCTGAGGAGTTGCGATTGCAGGC
        [A,C]
        CCTGTCACCATGCCCTGCTAATTTTTGCATTTTTTTGTTTGTTTTTTTTTTTTAGTAGAG
        ATGGGGTTTCACCATGTTGGCCAGGCTGGTCTCAAACTCCTAACCTCAAGTGATCACCCG
        CCTCAGCCTCCCAAAGTGCTGGGATTACAGGTGTGAGCCACCACACGTGGCTATGACCCT
        GATTTTGATTCATTCACTTTTTATAATTACCTTTTGATTAGATAAGTTAATTATTCTTGA
        ATTTGGCCATTTTATGCTTTGAGAAAGTAGTTAATCACAGTGGGTCAACAGTACAAACTT


40762   ATCCATCACCTCAAGCATTTATCCCTTGTGTTACAAACAATCCAATTACACTCTTAATTA
        TTTTTAAGTGTACAATTAAATTATTGAATATAGTTCAAAGACTTCTTCATTCATGACTAG
        CACCTAGGCTAAAAAAATTCAGACACCTGGGCTCCTGGGATCAATCACGCATACTGTGTC
        TCTTGTGCTCACTCCC
        [G,A]
        CTGTCTCTCTCTCTTTCTCTCGCTTCCTTTTTCCTCTCTCTCTGTGGTTTTCTAGGGTGG
        TGGCCTCAGGGAATTGGATTTCTTATATTATAGCTCAGGATTCCCAAGAGGGCTGTTTTT
        AATGTAGCCAAAGAAGTCTTGCAGCGTGACTTGTTTTATTCTATTCATTGAGGTAGTCAC
        AGAGGCCCGACCACAT


40936   TACAATTAAATTATTGAATATAGTTCAAAGACTTCTTCATTCATGACTAGCACCTAGGCT
        AAAAAAATTCAGACACCTGGGCTCCTGGGATCAATCACGCATACTGTGTCTCTTGTGCTC
        ACTCCCGCTGTCTCTCTCTCTTTCTCTCGCTTCCTTTTTCCTCTCTCTCTGTGGTTTTCT
        AGGGTGGTGGCCTCAGGGAATTGGATTTCTTATATTATAGCTCAGGATTCCCAAGAGGGC
        TGTTTTTAATGTAGCCAAAGAAGTCTTGCAGCGTGACTTGTTTTATTCTATTCATTGAGG
        [T,C]
        AGTCACAGAGGCCCGACCACATTCAGAGGAGGGACATACACTTGCTGGGACAAGTGTAAG
        AGAATTCATGATCATGTTTTAAAACCACTTTTATTAGTTTCCTATTGCTGCTGTAATAAA
        TTACCACAACTTAATGGCTTAAAAGCCACACAAATTTAATATCTTACAGTTCTGCAAATC
        AAAAGTCTGAAACGGATCTCACTGTGCTAAAATTAAGGTGTTCGTAGGGCATTCTGGAGG
        CTGTAGGAGAGAGTCTTGTTTTTTTGCCTTTTCTGGCTATTAAAAGCTGCCAGCATTCCTT


45998   TGTATATCAGTCAAAATATTGGGCAACTCTGATAAGTTTGTCCACTTAACATTGTACCAC
        TTAAGATGAATAGCATCTACCATTTCCGTCATTTGTAAATATATAGGAGGACATAATCAC
        ATAATCTTGAAGTAAAAGACAGTGCTTAAAACTGAATCAGTTAAGTTTTATGAAAAATAC
        TTCATATTGTACTTTTAAAAATATATATTTTTTAATTTCAATAGCTTTTGGGTTACAAGT
        GGTTTTGGTTACGTGGATGAATTCTATAATGGTGAAGTCTAAGATTTTACTGCAACTGTC
        [A,G]
        CCCAAGTAGTATATATTGTATCCAGCATATTGTCCTTTTTTTTTTTCTTTTTTTTTTTTCA
        TTTCACCATGGACTAATGAAAATTTTGTTAGGGACTGACATTAGGGCACCCTTGAGCTAC
        CTTGAGCTAAAGGAAATAACCCTTGAATTTTTTTCTGTTTGGCCTAGAGAATGTGGTTTG
```

FIGURE 3, page 25 of 27

```
        TTTTGTAACTGAATTCATGGGATTGTTAAGGTACAAGATTTTGCTTTAGTTTTATTTGTA
        CTAGGATTTTGCTATATTAATACAATGTGAAAAGAATCAAAAGTGTTAGAAATAAATGCA

47771   GAAGAGTAGAACATGAGGCTTTATTTAAAAGATTAGCAGAATTTAAGGAAAAGGTGACTT
        TGTTGAAGATTATAATGTGAAGACAAAGGAACGAGGATGGGAATAAATTTTGTATTCATG
        AGGCTTTGAAGAAATTGACTCTAGAGAGTATATTTTGGGTACTTTTGGGAAATGAAGTTG
        GATTAGTGAGAAGGAACAGATTATGAAAAGACAAGAAACCTGATTAATGTCAGGATGATT
        TTATATTTGAAG
        [T,C]
        TGGTCAGATTTATGGCAGTCCTGGCTTTGCCATTTTTAGTTTGATGACTTTGAGAAAGTT
        CCTTCTTGAAGTTTTAATTTTCTGTATATAAAAAGTAATAACACCTGGTGATCTGCTAGG
        TTTGTTTTGAGGATTATATGAGATAAAATGCATGCAAAACTGTTATAATAGTGCCTGGTA
        AAATAAGTGCCTAGTTTTAAAAACAAGTCTTTGTAAACTGCTTAGGACATGCCTGGTATA
        GGGTAGGTATGT

48117   GACTTTGAGAAAGTTCCTTCTTGAAGTTTTAATTTTCTGTATATAAAAAGTAATAACACC
        TGGTGATCTGCTAGGTTTGTTTTGAGGATTATATGAGATAAAATGCATGCAAAACTGTTA
        TAATAGTGCCTGGTAAAATAAGTGCCTAGTTTTAAAAACAAGTCTTTGTAAACTGCTTAG
        GACATGCCTGGTATAGGGTAGGTATGTAATACATAGTAGGTAGGATCTGTCTCCTTGCTA
        TTTTTAGGTAAAAAAACAAAAGGAAGAGCTTCAGCTTAATACAGTATGAACTGACGAGCC
        [C,T]
        TGGTAGGTTTTTGAGCAAAAGAGCAACACAGTAAAAGTAGTACTTAGGAAAGATTAACAA
        GGGAACATGGCTTATACAGTGGTAATGGGGCCTGGAGTCAAGGAGGTAAGATAAAATGGT
        ATTATAATTAAGGAATAGCCAGGCACGATGGCACATGCATGTAATGCCAGCTACTGGAGA
        GGCTGAGGTGGGAGGATCATGGGAGTCCAGGAGTTTGAGACCAGCCTGGGCAACTGAGTG
        AGACCCCAAATCCTAAAAAATACAAAGTAAAAAAGGAATAAAGTCATGAGGGCTTGGACT

54563   GCTTTGTCACCCAGGCTGGAGTGTGGTGGTATGATCATGGCTGACTGCAGCCCTGACCTT
        CCGGGCTCAAGTGATCTTTCCACCTCAGCCTCCCAATTACTTGGGACCACCAGCATGCTT
        GGCCGATTTTTTTTTTTTTTTTTTTTGTAGAAGCAAGGTTTCCCTATGTTGCCAAGGC
        TGGTCTTGAACTTTAGGGCTCATGTGATACTCCTGCCTCGGCCTCCCAAAGTGTTAGGAT
        TACAAGCCTGAGCCACCATGGCCGGCCAAAATATTTTCACTATAACAAATATCATATCTG
        [T,G]
        ATATACTCAGTTTTAATACTAACTCAAAGTAGAAACATAAAGCTGAATGACTATTTTATT
        TTCAGATTCTCTCCATTGAGTTTCCTTCTCCGTCTTGTGTGATCTCTGAACTTTTCTCCA
        TCTTTGCCACTTCTTGTCTAGCATTTTTTTTTTTATCAGCAGTTTCATTCAGATTTTTTTT
        TTAGTTCTTTCAACGGTGGAGTGGAAGTAGGCAGCAGGACAGAAGAACTTGAAGCAGAGC
        ACACTGGAGAGGAGAAATTAACAAAGCCTTTATGAATAAAACAACCCCCCAATATCAGTC

58735   TGGGTTATGCCCTGTTAACTCTTACATCATTAGTTTTTAGCCCAAAAGGAAACAGCAAAT
        AATGTTTTATATGAGCCACATTTTGCGTTGATTTTCCTTCCACTCTGTAAAATTACTAAA
        GCAGCACTCTGACTTTATTATGCTCAAATCGCTCTTCTCCATTAATGTGTGTTTCTCCAT
        CTTTTTAGGGTTTTTTACTTTATAAATACAGAGATTACTGTGTAAAATTCTAAATTTGCCAC
        TGGGTCGTTATACATTTGTAACCTTCCTCACAGTATATTTTGTGATTTGGCAGAGTTTAC
        [C,T]
        AATATAGATGATACTAACTGAAATTAATCATTCTGTATAATTGGATAGAAAAGCATGAGT
        AAGAATTCAATTGGTATTATATTTAATTAATTGCCAAGATTTTCACATTTCCTGACTACA
        ACAATAAAATCAAATGAATTGATGGCTTAAAAAAAAGAAATCTCAAATGTTTAGTCAATG
        AAGAACATCTATTGAATGAGTGAATGTTCATTATATATAGTGCATTTTCTGAGCTTTTTT
        GGAGGGGGAAGTTGCTCCCATGCTCTGAGAACTTTTAAGGATCGATACATTATTTTTAAC

59643   GTTTATATTGCCACATTAATTTCCATTATAAAACCAGTAACCATAGTTTTGTTTTAATTA
        GCAATCTAATTATTTTCATGTATCCTCATTATGAGAATTTATGTCCATCACTTTGCTTGA
        TGTGATAACAGTGACATGCTAAATGAGAAACAATTGTTATTTAGAAAAAAAATGCACAAAG
        TGAAAGTCCTTTTAATCCCTAATCATAAATACATTTTATTAGCTTACTTTAAGAAGTGGC
        AGTCACAGCTCCTGAACATTAGGGAGTGTTTCTTTTGGTCAGCATTATTTATTTAGTGCA
        [C,A]
        ATTGCCTTTAATTTTAATTTGAAATTATAGTAAAATCCACGGGAGTTTTTAAGTCTCCTC
        ACAGCCTTTTGCTACCTTTTCACCAAGGTAGATCCAGATGATAACTGCTGTGTTGTGACA
        TCATAGAAATTAGAAAAATATTTTCCTCTGAGGAAAGAACATTGTAAATGAAACTCTACA
        TATCAGAGGTCTATAGCTATGTATCAATATTAAGTTTCTTTTGTACTTTGCTTTGTAGTC
        ATCTTCATTCCAAACTTTCATAATTATTATTTTTACTTTAAAAAGAAAAATAACCCACCA
```

<div align="center">

FIGURE 3, page 26 of 27

</div>

61638   AAAAAAAGGAAAACATTGATAAGTGTCCTAGAAACTTGGATTCTTTTATAGATTTGTTCT
TGGGGCTCTGATGTTTGGGATTGACGTTCTGTGCTGACCATTTTATATGCATTTTATCTT
AATAGTATGTGCTTTCATGAAGATTCTGATACAAGTGGGCAATCCTTAAATTATCTTTGA
AAAATTGGTTAATTTTGGTTAAAAAAGGGAAAGTGGCTGGGTGCAGTGGCTCACGCCTGT
AATCCCCAGCACTTTGGGAGGCCGGGACGGGTGGATCACAAGGTCAGGAGTTGAAGCCCA
[G,T]
TCTGGCCAACATGGTGAAACCCTGTCTCTACTGAAAATAATTGGGGCATGGTGGCACATG
CCTGTAATCCCAGCTACTTGGGAAGCTGAGGCAGGAGAATTGCTTGAACCGGGGACCCAG
GAGGCGGAGGTTGCAGTGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGCTACAGAGC
GAGACTCTGTCTCAAAAAATAAATAAATAAATAAATGAAAAAGAGAAAATATTGAGAGGA
TTTGGTCATCATTTTACTGCTCTCTTCATGTGATGGAAATCAATTTTCCTTCTCAAATGG

63291   GAGATGTACTGTGATTTTACTGAGGTTTCATCACAAGAAGGGAGTGTTTCTTGTGCCATT
AACCATGTAGTTTGTACCATCACTAAATGCTTGGAACAGTACACATGCACCACAACAAAG
GCTCATCAAACAGGTAAAGTCTCGAAGGAAGCGAGAACGAAATCTCTCATTGTGTGCCGT
GTGGCTCAAAACCGAAAACAATGAAGCTTGGTTTTAAAGGATAAAGTTTTCTTTTTTGTT
TTCCTCTCAGACTTTATGGATAATGTGACCGGGTCTTATGCAAATTTTCTATTTCTAAAA
[G,C]
TACTACTATGATATACAAGTGCTGTTGAGCATAATTAAATAAAATGCTGCTGCTTTGACA
GTAAAGAGAAGGAAGTATTCTGATTAGCTGTATCTGGTATTAATTGCATGTTAAAACACT
GGAATTTTTAAAATTGAAATTAGATCAGTCATTCTTTTCTTTTCTCAAGATATCTCATGG
CTGACACTGAAGAAGAAATGTAATTCATAACTTGCACTAAATGTATATTTTTTTTCTTAA
AAATTTACCATTCTTATTTATATTTTTATGGATTAAAATTTATAAAATACAGATCAGTTA

63463   TGTGCCGTGTGGCTCAAAACCGAAAACAATGAAGCTTGGTTTTAAAGGATAAAGTTTTCT
TTTTTGTTTTCCTCTCAGACTTTATGGATAATGTGACCGGGTCTTATGCAAATTTTCTAT
TTCTAAAACTACTACTATGATATACAAGTGCTGTTGAGCATAATTAAATAAAATGCTGCT
GCTTTGACAGTAAAGAGAAGGAAGTATTCTGATTAGCTGTATCTGGTATTAATTGCATGT
TAAAACACTGGAATTTTTAAAATTGAAATTAGATCAGTCATTCTTTTCTTTTCTCAAGAT
[A,G]
TCTCATGGCTGACACTGAAGAAGAAATGTAATTCATAACTTGCACTAAATGTATATTTTT
TTTCTTAAAAATTTACCATTCTTATTTATATTTTTATGGATTAAAATTTATAAAATACAG
ATCAGTTAATATTGCACTTAAGTAATTTTACCTTTTTAATGTGATTTTTATAGAATAATT
CAGACTTACAAATACAGAGATATGAACAAAGTTTACAGTGGGAACAAAGGTTTAAAAAAA
GGTTGTGGTTCTCTCTCTGTGATCCAGTGTGCACATAAACCTTTCTCTGATCTTTCACTG

63636   TGCTGCTGCTTTGACAGTAAAGAGAAGGAAGTATTCTGATTAGCTGTATCTGGTATTAAT
TGCATGTTAAAACACTGGAATTTTTAAAATTGAAATTAGATCAGTCATTCTTTTCTTTTC
TCAAGATATCTCATGGCTGACACTGAAGAAGAAATGTAATTCATAACTTGCACTAAATGT
ATATTTTTTTTCTTAAAAATTTACCATTCTTATTTATATTTTTATGGATTAAAATTTATA
AAATACAGATCAGTTAATATTGCACTTAAGTAATTTTACCTTTTTAATGTGATTTTTATA
[G,A]
AATAATTCAGACTTACAAATACAGAGATATGAACAAAGTTTACAGTGGGAACAAAGGTTT
AAAAAAGGTTGTGGTTCTCTCTCTGTGATCCAGTGTGCACATAAACCTTTCTCTGATCT
TTCACTGCCATCCTCTGGATTATGTCTTCTGACCTGTCCATTTTGACCCATTAACTGGAA
AGTTGAAAAACTACATTAACTGGAAAGTTGAAAAACTACATTACTTTGGAGAATAAAACC
GAAAGTTCGTGTATACCTTCTTAAAAAAAAAATCAAACCAAAAATGTGAAAACAATAGAA

63998   AAAAAAGGTTGTGGTTCTCTCTCTGTGATCCAGTGTGCACATAAACCTTTCTCTGATCTT
TCACTGCCATCCTCTGGATTATGTCTTCTGACCTGTCCATTTTGACCCATTAACTGGAAA
GTTGAAAAACTACATTAACTGGAAAGTTGAAAAACTACATTACTTTGGAGAATAAAACCG
AAAGTTCGTGTATACCTTCTTAAAAAAAAAAATCAAACCAAAAATGTGAAAACAATAGAAT
TGCAAAGATAGCAGTTAAAATTTTAATCTGAAAATAACCTTTGAATCTCGGGCTAGGTTA
[T,C]
GTCCATATTTGAAGTGGTCAGTGATGGTTTGAACATTTTTTGCAGGATGAGTTAAAATGC
ACTGGATTATATTTGGGATTTTTGTTTTTGGAATTGTCTGTTTTAATCACAGCCTTAATT
CACAATTGGCAAAGGCAGTTTACTCAAAGGACTGGGCTAAATATTCTGTAATTATGCATT
TTTGATAGGAAAATGAAATTTTTGCAAACAGACATTTTCTTTTTTTTTGGCTGGAGTGCA
GTGGGGCATGGTCTTGGCTCACTGCAGCGTTGACCACCTGGGCTCAAGTGATACTCCCGC

FIGURE 3, page 27 of 27